JAN 20 1928

The Journal of EDUCATIONAL PSYCHOLOGY

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

BOARD OF EDITORS

HAROLD RUGG, *Chairman.*

J. CARLETON BELL — ARTHUR I. GATES — RUDOLF PINTNER — L. M. TERMAN

FRANK N. FREEMAN — V. A. C. HENMON — BEARDSLEY RUML — E. L. THORNDIKE

VOL. XIX
No. 1

JANUARY, 1928

$4.00 a year.
60c per copy.

CONTENTS

OF MENTAL TESTING—A PRAGMATIC THEORY. *Charles S. Slocombe* 1

THE RELATIVE INFLUENCE OF FATIGUE ON PUPILS WITH DIFFERENT LEVELS OF MENTALITY. *Grayson N. Kefauver* 25

A STUDY OF CAUSES OF RETARDATION AMONG MEXICAN CHILDREN IN A SMALL PUBLIC SCHOOL SYSTEM IN ARIZONA. *O. K. Garretson* 31

THE MEASUREMENT OF CHILDREN'S CONCEPTS. *Laurance F. Shaffer* 41

A MODIFICATION OF TOLLEY AND EZEKIEL'S METHOD OF HANDLING MULTIPLE CORRELATION PROBLEMS. *Henry E. Garrett* . . . 45

LEARNING CURVES OF BOYS OF IQ's 76–148. *Frank T. Wilson* 50

A PEARSON-*r* FORM FOR USE WITH CALCULATING MACHINES. *Luton Ackerson* 58

A SIMPLE APPARATUS FOR THE CLASS DEMONSTRATION OF THE CONDITIONED RESPONSE. *George Humphrey* 61

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES. 64

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION. 68

Published Monthly Except June to August by

Warwick & York, Inc.,

York, Pa. Baltimore, Md.

Entered as Second Class matter November 15, 1921, at the Post Office at York, Pennsylvania under the Act of March 3, 1879.

The Journal of Educational Psychology

is published monthly, by Warwick & York, Inc., Baltimore, Md., and York, Pa. Title page and index are bound in December number.

Manuscripts for publication, books or other materials for review, and news items should be addressed to Harold Rugg, Chairman of the Editorial Board, 400 W. 118th Street, New York City.

The price of the Journal is $4.00 a year, payable in advance. Foreign postage is 30 cents extra. Single current issues are 60 cents each, and less than a full year's subscription is charged at single issue rate for every month ordered. Back volumes $5.00 each; single back issues 90 cents each.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which change is to take effect. No claim for nonreceipt of an issue will be entertained unless made within two weeks after receipt of the next succeeding issue.

Warwick & York, Inc.,

Baltimore, Md.

TEACHERS' HEADQUARTERS

European

Plan

Entirely

Fire-Proof

Centrally

Located

EDWARD DAVIS

Manager

HOTEL RENNERT

BALTIMORE, MARYLAND

UNIVERSITY OF VIRGINIA SUMMER QUARTER

EDWIN ANDERSON ALDERMAN
Ph.D., D.C.L., LL.D., President
CHARLES GILMORE MAPHIS
Ped.D., LL.D., Dean

First Term, June 18–July 27
Second Term, July 30, Aug. 31
Institute of Public Affairs, August 6 to 18

The Summer Quarter is an integral part of the University year.

Degrees are conferred upon men and women for summer work.

The Master's Degree may be obtained by properly qualified students in three summer quarters.

It offers opportunities unexcelled in the South and makes a strong appeal to teachers seeking broader scholarship and training and wider social contacts, and to college students desiring to complete degree requirements.

The Institute of Public Affairs attracted nation-wide attention. It offers excellent opportunities for students of government and the social sciences.

Attendance last quarter, 2,167 from 35 states, and 6 foreign countries.

The most beautiful and unique campus in America.

Accommodations at reasonable rates. Tuition for non-Virginia students, $25.00 per term.

Entertainments, music festival, excursions.

Reduced railroad rates.

For illustrated folder and full announcement write to

Secretary, SUMMER QUARTER
University, Virginia
Box 149

SCHOOL AND SOCIETY

A weekly journal covering the field of education in relation to the problems of American democracy.

The objects of SCHOOL AND SOCIETY are the advancement of education as a science and the adjustment of our lower and higher schools to the needs of modern life.

It encourages discussion of all problems from every point of view; it pays special attention to foreign affairs; it publishes accounts of educational events of the past week and research on fundamental problems.

Subscription $5.00—Single Numbers 15 Cents

[*SCHOOL AND SOCIETY will be sent to new subscribers for thirteen weeks for one dollar, which may be sent at the risk of the publisher.*]

THE SCIENCE PRESS

Grand Central Terminal - - New York, N. Y.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XIX | January, 1928 | Number 1

OF MENTAL TESTING—A PRAGMATIC THEORY

CHARLES S. SLOCOMBE

Personnel Research Federation, New York City

Introductory.—Mental testing, both of Intelligence and of Special Abilities, while it has made some considerable progress during the past quarter of a century, seems now to have almost ceased to develop. While the use of tests is increasing, their value as measures of intelligence or of abilities has not improved to an extent commensurate with the amount of research and experimentation to which they have been subjected. The difficulty that has greatly troubled all test builders is the absence of some reputable theory as to the nature of intelligence. It is, of course, absurd for any psychologist to attempt the construction of tests of intelligence, abilities, or what not, without carefully and clearly defining what is proposed to be tested; for it is impossible under that circumstance to evaluate tests, and in fact their construction must be largely a chance hit-or-miss proceeding.

The majority of theories which have been put forward have been very limited in their application (in many cases it is difficult to see any application) because they have not been ultimate in their analyses of mental processes. Insofar as they have been stated in physiological terms they have been of no assistance whatever.

In view of this situation, there appears to be place for a theory of intelligence, abilities, and testing which does not suffer from such limitations. Spearman's theory is free in this respect. It is capable of direct application in the construction of tests—determining the selection of items; it is unique in that the psychological theory is closely related to and integrated with numerical results obtained, and the statistical treatment and interpretation of those results; and, furthermore, it links the test movement to general psychology, and does not leave it as a detached method, bearing little or no relation to other experimental work. The paper here presented attempts an exposi-

tion of the theory, its development, and the meaning it has in statistical terms.[1] No attempt has been made, however, to offer mathematical proofs, and in fact, the statistical treatment has been maintained at a minimum.

Part I. Theory of Intelligence

Cognitive Processes.—Spearman[2] enunciated the following three principles, which he claimed include all mental processes capable of generating new items in the field of cognition.

1. Any lived experience tends to evoke immediately a knowing of its characters and experiencer.

2. The mentally presenting of any two characters (simple or complex) tends to evoke immediately a knowing of the relations between them.

3. The presenting of any character together with any relation tends to evoke immediately a knowing of the correlative character.

(The term "immediately" is used in a special sense—meaning the absence of any intervening stages or processes.)

These three principles he calls qualitative, noegenic, that is "self evident" and "generative of new items," and claims that all mental processes, such as reasoning, synthesis, generalization, etc., which result in the formation of new items reduce on analysis to one or more of these principles.

To these he adds five quantitative principles, cognitive control (attention), fatigue, retentivity, constancy of output, and primordial potency which do not regulate the quality, but only the quantity of cognition; deriving from these five are three anogenic principles, reproduction, disparition, and clearness-variation.

This analysis of mental processes, which cuts across the traditional classification, is of very considerable importance for psychology generally. It has been criticised, and not more widely acknowledged because of the formality of the principles as stated. The criticism is really absurd, for insofar as they define universal tendencies, they cannot relate to specific material. The laws of all sciences are formal. Were Newton's laws neglected because of their formality? "To every reaction there is an equal and opposite reaction." Is this state-

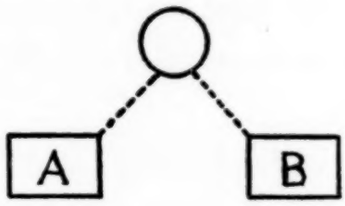
[1] The author is indebted to Professors Spearman and Woodworth for many helpful suggestions regarding method of presentation.

[2] Spearman, C.: "The Nature of Intelligence and Principles of Cognition." 1923.

ment more meaningful than "Any lived experience tends to evoke immediately a knowing of its characters and experiencer?" Is it more self-evident? More logical? Capable of wider application? It seems as if psychologists have been for so long accustomed to amusing themselves with "a string of raw facts, a little gossip and wrangle about opinions"[1] (which they regard as interpretations of those facts) that they cannot conceive their subject as a science—an organized body of knowledge, having postulates, laws, and principles, stated in clearly defined and ultimate terms.

The value of the principles will be considered here solely in the realm of mental testing, into which they can introduce some logical *raison d'être.*

Second Principle.—Of the principles only the processes included in the last two classed as noegenic are important for this purpose. Their meaning may be clarified by examples. The second is capable of representation thus:



A and B are two presented characters—fundaments—which according to the first principle have been apprehended. The dotted line represents the relation which is educed between A and B, and constitutes the new item which is generated. Thus, on the table near this paper is a red pencil, and beside it a match with a red tip. These two objects may be regarded as the fundaments A and B. A relation is educed between them—one of difference of redness. Other relations may be evoked, difference in size, in dullness of red, and so on. Or perhaps each object is related to the green table top, as different from it, and similar to each other, in that both are red to some extent. All these relations are new items generated, evoked or educed as a result of the presentation of the things or fundaments. This seeing of a relation is "contingent," that is it *may*—not *must*—happen. On the table nearby has been an ink-pot and lampstand, but these things have not until just now been related. There are many other things on the table and in the room between which there has not and probably never will be educed relations.

[1] James, W.: "Psychology, Briefer Course." 1892, p. 468.

In a mental test however, fundaments are presented together, and the subject is required to educe a relation between them. In a good test the two fundaments chosen, though their significance be known, should not ever before in the experience of the testee have been related. In view also of the required comparability of responses and scores, only one relation should be capable of eduction by the particular group of subjects being tested.

Examples.—In a same-or-opposites test the fundaments might be, *Top—Bottom; Evil—Wicked.* In a test of discrimination in music they might be two tones. An analogies test provides a field in which a much wider selection of relations is possible.

Bat:Ball::Hammer:*Handle, Nail, Boy, Axe*

In this test the subject educes a relation between the first two fundaments

Bat—is used for hitting—Ball

He may then relate the third fundament to the first choice.

Hammer—has—*Handle*

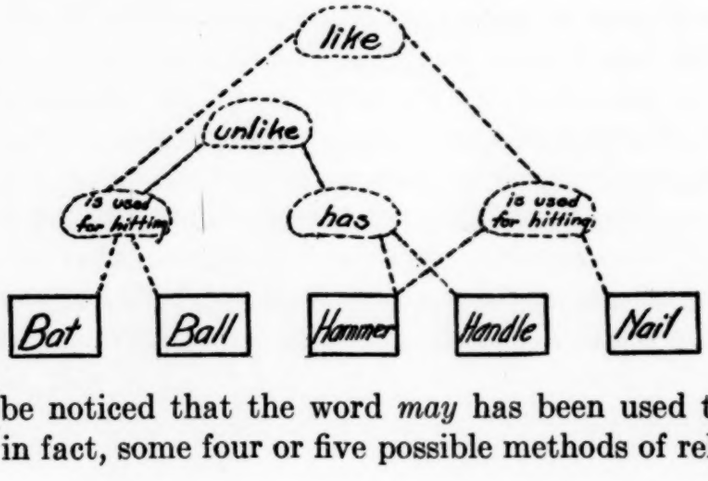
relate the two relations,

"is used for Hitting" and "has,"

and see that they are different, instead of being similar as required. He may then relate the third fundament to the second choice,

Hammer—is used for hitting—*Nail*

relate this relation to the one between Bat and Ball, see that it is a like one as required, and select it as his response. The whole process may be schematically represented.



It will be noticed that the word *may* has been used throughout. There are, in fact, some four or five possible methods of relating these

fundaments to obtain the correct response. The method described here has been used because it illustrates the eduction of relations. The proceedings consist altogether in

Eduction of a relation between 2 fundaments, Bat and Ball
Eduction of a relation between 2 fundaments, *Hammer* and *Handle*
Eduction of a relation between 2 relations, "is used for," "has"
Relation of this relation to the required relation of similarity
Eduction of a relation between 2 fundaments, *Hammer* and *Nail*
Eduction of a relation between 2 relations, "is used for," is used for
Relation of this relation to required relation.

On ultimate analysis there has then been educed

(*a*) 3 relations of 1st order—between fundaments,
(*b*) 2 relations of 2nd order—between relations,
(*c*) 2 relations of 3rd order—between relations (from (*b*))

All this relating takes place in a very short interval of time. Possibly some will deny that it takes place at all. If, however, the same test is given to a mental defective, or a very difficult set of fundaments is given to an adult so that the processes are slowed down, the actual progression of steps can be demonstrated. By inference[1] then one is justified in stating that it takes place in all cases, but perhaps very very quickly.

As a difficult example, the following illustrates well the processes of eduction involved in answering test questions.

In the first line the letters have been underlined according to a certain scheme. Underline the letters in the second line according to the same scheme.

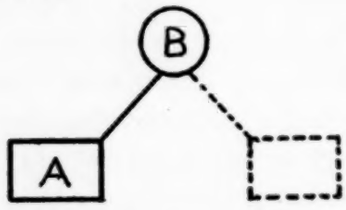
mnpq	ponm	qmpo	nqol
trsu	sruv	twsv	rwvs

As illustrating the diversity of relations which may be educed by different individuals, and the great care which must be taken to ensure that only one relation can possibly be correctly educed, the response of one subject may be mentioned. One child, on being tested with the above analogy chose *Handle* as his response, justifying himself on the ground that Bat and Ball began with the same letter, and that this

[1] Spearman, C.: New Psychology of Shape. *British Journal of Psychology*, Vol. XIV, p. 211.

being the required relation, chose *Handle* and *Hammer*, as also beginning with the same letter.

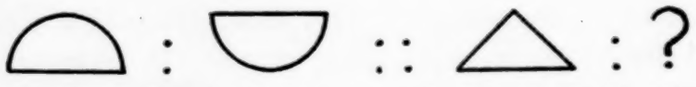
Third Principle.—The third principle may be represented by a somewhat similar diagram.



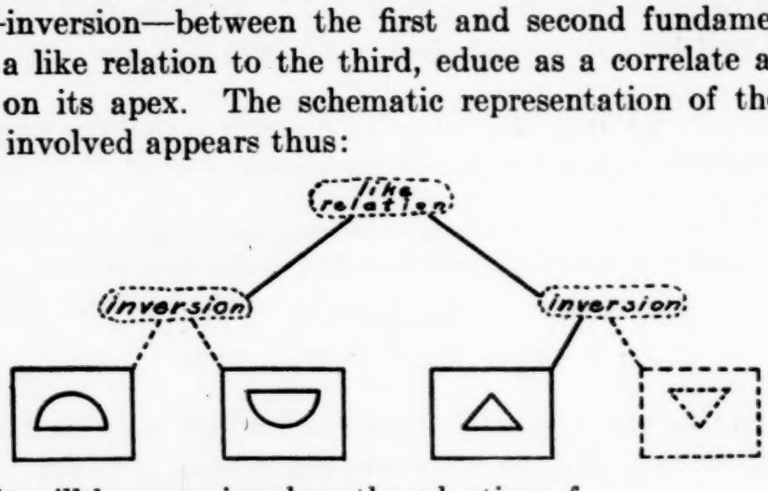
In this figure the continuous line represents the fundament and the dotted line a correlate which is educed.

Example.—Give the opposite of *High*. *High* is the fundament, oppositeness the given relation, and *Low*, if educed would be the correlate.

Again the process may involve higher levels, as in solving the analogies test.



Here the subject has to draw a fourth figure which bears the same relation to the third as the second does to the first. He might educe a relation—inversion—between the first and second fundaments, and applying a like relation to the third, educe as a correlate a triangle standing on its apex. The schematic representation of the mental processes involved appears thus:



This it will be seen, involves the eduction of

1 relation of 1st order
1 correlate eduction of 2nd order
1 correlate eduction of 1st order

As with the other two noegenic principles the eduction of the correlate is contingent. It tends to happen, but may not.

This third principle is similar in some respects to the process which may be termed Reproduction or Recall. Quite often tests which are regarded as of memory—particularly as of logical memory—really involve the eduction of correlates from the context, especially when memory fails. On the other hand in very many intelligence tests, an analysis of errors reveals the fact that the majority of these are caused by the use of reproduction (generally association) when the eduction of correlates is required. Also test items which are found to be unsatisfactory as assisting in the measurement of intelligence, are generally found to depend on reproduction. The main difference between these processes is evidently the fact that only the latter is capable of generating new items, which are not limited by previous experience. It appears that this process of eduction is really what takes place in what are generally termed invention and imagination.

Though Analogies and Opposites, as types of test, have been mainly used in illustrating the principles, all other forms are capable of analysis in these terms; in fact, all mental processes whatsoever may be reduced to the principles, either noetic or anoetic.

Definition of Intelligence.—"The word intelligence must be taken to include all the processes derived from all three principles." So that to test intelligence is to test the ability to perform the mental operations or processes involved in apprehending, educing relations and educing correlates. It is certain, of course, that the cognitive processes involved are not limited to these three, but that reproduction, shift of attention, disparition, etc. also occur, and that a great variety of other factors enter into and constitute the "total situation" of testing.

The unavoidable presence of these extraneous factors has occasioned the greatest bar to progress in the development of tests, and has been the main argument of opponents in their criticisms of the movement. Mental testers have been hard put to it to justify their existence, and have in general had to fall back on an empirical "somehow it works." Spearman's contribution to the movement has been to enunciate and demonstrate the working of a theory which is statistically and logically sound, and which is immediately applicable in the building and evaluation of tests. It is furthermore capable of development. The enunciation of a theory obviously cannot solve all the difficulties of testing, but it should form the basis of systematic research.

Classification of Relations and Fundaments.—It will have been evident from the above illustrations that the mental processes defined in the principles considered cannot take place *in vacuo*. Something

must be apprehended; a particular relation must be educed between particular fundaments: The knowing of a correlative character results from the presentation of a particular relation and a certain fundament. The range of relations and fundaments is almost infinite, but classification is possible, though exceedingly difficult. The fundaments, including all material of the first principle may be classified as

(*a*) Sensory experiences, having quality (subclassified as sight, sound, taste, smell, and tactile sensations), intensity, spatiality, and temporality.
(*b*) Affective experiences.
(*c*) Conative experiences (impulses, desires, and the like).
(*d*) Cognitive experiences.
(*e*) Relations which have been educed. (It is possible as in the analogies test above, to educe relations between relations—these are spoken of as being of the 2nd and 3rd order or level.)
(*f*) Correlates which have been educed.
(*g*) The ego.

Relations may be classified as

Real	Ideal
Attribution	Evidence
Identity	Conjunction
Time	Intermixture
Space	
Cause	
Objectivity	
Constitution	

Analysis of Situation.—A mental process, under the second principle may be analyzed into its constituents in order to ascertain the conditions which make it contingent. (Similar analysis is possible with other principles.) It would appear that a person's ability to educe a relation depends upon the following factors:

(*a*) His ability to see relations of any kind.
(*b*) His familiarity with each of the fundaments.
(*c*) His familiarity with the required relation.
(*d*) His willingness to use his ability.
(*e*) His general condition, fatigue, emotionality, etc.

These factors enter into the total situation, and constitute a circumstance in which a relation is or is not educed. (The relationship between these factors and the situation is not conjunction but constitution.)

The ability to educe relations and correlates correctly is the function which tests attempt to measure. For this reason it should be the crucial factor. A person should not fail in a test item because of his lack of knowledge of the relations and fundaments, but because, having that knowledge, he cannot use it. That is to say, a test should be constructed of such material that the results are not influenced by the nature of the material. This assumes what is called a "common background" for all individuals tested. The inclusion of many items permits differences in background—of a given type. But if backgrounds differ then must items be included which are representative of them all. The difficulty which is involved in the comparison of the intelligence of individuals from groups differing socially, racially, etc. is due to the fact that all tests used have included group factors. For example, with the London Revision of the Binet scale, Burt found that the average child from a poor district was lower in intelligence than one from a better district. The interpretation is open to question, for there are two possibilities: (*a*) The poorer children are on the average less intelligent; (*b*) the test results are influenced by environmental conditions. In order to eliminate the second possibility it would be necessary to use a test which places the richer children at a disadvantage comparable with that under which poor children are working when tested with the London Revision. Pintner has discussed this difficulty, having arrived at its consideration from behavioristic view of intelligence.[1] An unwillingness to concede the necessity for an analytical and theoretical consideration of tests has resulted in the testing of different racial groups, (as Nordics, emigrants, and so on) and the arrival at conclusions which are unjustified and probably unjust.

With reference to determinant (*d*) little work has been done to investigate its influence. Courtis[2] has considered the subject in scholastic tests, and found the influence considerable. Generally it is assumed that testees do put forth a maximum of effort, but this is an assump-

[1] Pintner, R.: An Empirical View of Intelligence. *Journal Educational Psychology*, 1926, p. 608. Pintner considers that any "reaction" might be used in testing intelligence. "There is nothing inherent in a reaction that makes it rather than another reaction, a test of intelligence." The view here expressed is that while intelligence enters into *all* "reactions" (with great stress on *all*) to some extent, nevertheless all reactions are *not* suitable for use as tests of intelligence. For if a test is of intelligence, intelligence, and not any other factor must be crucial.

[2] Courtis, S. A.: "Why Children Succeed." 1923, p. 130.

tion. It may be, however, that the amount of effort put forth is specific to the item, or test, or occasion. In the wider problem of the testing of very young children, delinquents, and the insane, the influence of this factor must be very large, and results of such testing should be very carefully examined, in view of this fact.

It is expected that the effect of determinant (*e*) is eliminated by testing on more than one occasion. Insofar as it is not, it must be regarded as an error in measurement—which on account of the trouble and expense of elimination must be retained.

Part II. Application to Tests of Intelligence

What Is Measured.—Suppose the following to be items in a subtest.

Give the opposite of:

1. Mountain.
2. Dirty.

. . . . .

19. Obsequious.
20. Denominator.

An individual's ability to give the opposite to Mountain might be said to depend upon (*a*) his general ability, (*b*) his knowledge of what a mountain is, (*c*) his understanding of the relationship of oppositeness, (*d*) his knowledge of physiographic features, plains, valleys, hills, etc., (*e*) his condition at the time of test, (*f*) accidental happenings, distractions, etc. Similar analyses might be made in reference to other items. His knowledge of physiographic features is independent of his knowledge of "obsequious" or "denominator" and other words. That is to say, any Grade IV boy may know any one of these, but from that fact it cannot be said whether he will know any other. The factors (*b*) and (*d*) are then specific to each item, and independent of all others. So the ability in this subtest may be represented thus:

1. $g + s_1 + s_2 + s_0 + s_1 + e$
2. $g + s_3 + s_4 + s_0 + s_2 + e$

. . . . .

. . . . .

19. $g + s_6 + s_7 + s_0 + s_4 + e$
20. $g + s_8 + s_9 + s_0 + s_5 + e$

g stands for general ability; s_0 for understanding of the relation of oppositeness; s_1, s_2, . . . for specific factors; and e for accidental errors.

It will be seen that g is common to all; that s_0 is common; and that e is common. $s_1, s_2, \ldots s_n$ and $e_1, e_2, \ldots e_n$ are independent and occur only once each. If s_1 be knowledge of mountains and s_9 be knowledge of fractions, then it is a matter of chance whether a subject knows either, neither or both. If a large number of such s's be included they will mutually cancel on addition of the whole, for the number that are known will be about equal to the number that are not. If therefore the items are added the pooled result will be $(g + s_0 + e)$.

Suppose a test to be made up of a number of subtests, Analogies, Opposites, Completion, and so on. Then the scores on all these subtests might be represented thus:

$$\begin{aligned} \text{Score on Opposites} &= g + s_0 + e \\ \text{Analogies} &= g + s_c + e \\ \text{Completion} &= g + s_a + e \\ &\ldots\ldots \\ \hline \text{Total} &= g + e \end{aligned}$$

The s's being independent cancel out, if there is a sufficient number of subtests, and there is left only a measure of g and an error.

Illustration has been given of the measurement of the general ability to perform the mental operations included within the categories of the noegenic principles, and has consisted in the consideration of a test in which all the other factors entering into the situation have, as far as possible, been eliminated. All the various types of relations and fundaments classified above would require inclusion if the elimination is to be as complete as possible.

The Nature of Error.—The nature of the error, which has perhaps been erroneously represented (to avoid too many subscripts) as the same throughout, requires some little explanation. A wide variety of factors may contribute to it, some of them correlated, and therefore remaining as an error in the score. Among those which remain would be the physical and mental condition at the hour of testing, the general atmosphere of the test situation, and the verbal medium in one language, in which the test is set. It will be seen that the first two of these types of error are specific to the occasion of testing. They could be eliminated by testing on successive occasions, when like all other specific factors they would cancel out. The last type of error, verbality, can also be eliminated by setting a test in different media: Verbal, pictorial, concrete, motor and so on, when the factors specific

to each would tend on the whole to cancel. The difficulties and inaccuracies introduced by such means, however, are of such a nature that error of another kind is introduced, so that it is advisable under present conditions to remain content with the degree of approximation achieved by verbal tests, except in the special circumstances discussed above.

The Correlation of Abilities.—The logical consideration of intelligence, and the consequent nature of tests designed to measure that function has led naturally to the General Factor Theory enunciated by Spearman, as a result of his statistical analyses of test material and scores. Having indicated what a test attempts to measure, and the difficulties involved it is now necessary to demonstrate the fact that the statistical findings are in accord with that indication, and the extent to which the difficulties have and can be overcome.

It has been found that practically all abilities show some degree of correlation. Even such diverse activities as solving puzzles, discriminating between musical tones, tapping, arranging wooden blocks, and so on. The degree of correlation varies considerably, and between some abilities it may be very low. For example, the correlation between tapping and an analogies test with College freshmen might not show at all by ordinary testing methods and treatment of results. With less developed individuals however there is significant correlation. But with increasing age or maturity the correlation reduces until it seems to be absent. This is due to the fact that the necessary muscular control becomes such that the tapping becomes mechanical; that is, the amount of cognitive control is reduced. There seems some justification, for the inference, however, that though the correlation has reduced almost to the vanishing point, it has not altogether ceased to exist, and that by much more careful experimentation and treatment of results some measure of correlation could still be shown to exist.

Spearman's General Factor theory accounts for this correlation of abilities. The theory states that "each ability is capable of resolution into two factors, (*a*) a general one, entering into *all* abilities of a person, termed *g*, and (*b*) a factor (or number of factors) specific to the particular ability concerned, termed *s*." Consider the following table of intercorrelations of nine Intelligence tests.[1]

---

[1] This table is extracted from that given by Miller: Variation and Significance of IQ's. *Journal Educational Psychology*, 1924. It has been discussed in another connection by the author of this paper in the same Journal, 1926, p. 600.

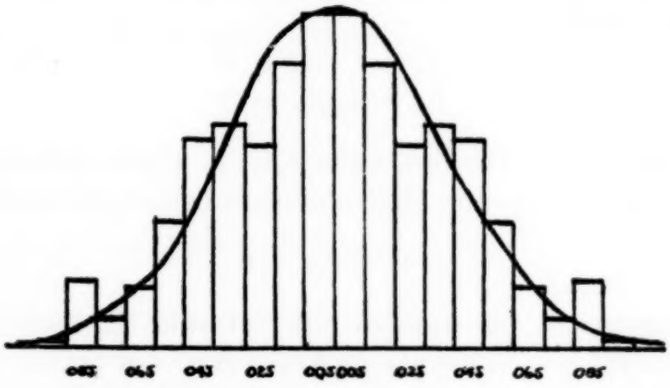
TABLE I

	1	2	3	4	5	6	7	8
1	..	75	70	89	82	87	74	86
2	75	..	70	77	76	79	70	76
3	70	70	..	77	77	74	70	76
4	89	77	77	..	90	90	77	83
5	82	76	77	90	..	83	75	82
6	87	79	74	90	83	..	77	86
7	74	70	70	77	75	77	..	78
8	86	76	76	83	82	86	78	

Each test correlates with every other test. Why? There are two possible explanations. (1) That there is something common to each and all which causes the correlation. (2) That between any two or

FIGURE I

Distribution of Tetrad Differences
(calculated in Table I)



more tests there is something common, but that there is nothing common to all tests. The former explanation is that represented by the General Factor theory of Spearman. The latter is the Sampling theory held by Thomson.[1] The correctness of either explanation

[1] Brown and Thomson: "The Essentials of Mental Measurement." 1925. Thomson regards the criterion put forward by Spearman in 1910 as inadequate to *prove* his theory, and derives his conviction of the truth of his Sampling theory from certain dice-tossing experiments he made, in which he satisfied Spearman's criterion, without, as he thought having any general factor present. A careful examination of Thomson's procedure however reveals the fact that in all x's between which there is significant correlation, there *is* a general factor (in fact two), and leads to the conclusion that had Thomson's results *not* conformed to the criterion, he would have disproved Spearman's theory.

depends upon certain criteria, on which there have been many mathematical papers published.

The Criterion in the General Factor Theory.—Spearman and Maxwell Garnett[1] have proven that the equation

$$\frac{r_{ac}}{r_{ad}} = \frac{r_{bc}}{r_{bd}} \tag{1}$$

holds "when and only when, each ability is resolvable into two factors g and s, where g is common to all abilities, and s is uncorrelated with either g or any other s." This relation may be derived in the following manner, as showing its significance in terms of more familiar formulas.

If g is the only factor common to all tests then the correlation between any two will be zero, if g is eliminated by partial correlation. The formula for this, by the ordinary method is

$$r_{ac \cdot g} = \frac{r_{ac} - r_{ag}r_{cg}}{\sqrt{1 - r_{ag}^2}\sqrt{1 - r_{cg}^2}} \tag{2}$$

But r_{ag} and r_{cg} are the correlations of each of these tests with g, and as $r_{ac \cdot g} = 0$, therefore

$$r_{ac} - r_{ag}r_{cg} = 0$$

Similarly

$$r_{bc} - r_{bg}r_{cg} = 0$$

Whence, by simple algebra, equation (1) above follows.

Therefore it may be said that equation (1), or its equivalent form

$$r_{ac}r_{bd} - r_{bc}r_{ad} = 0 \tag{3}$$

is a formula very easy of application (formula (2) could be used) to determine in a table of correlations whether or not two factors only are present. Its usual application is to find whether a test has been constructed so that all specific factors cancel out. It may also be used to determine whether any number of tests or subtests have factors other than g common to them, and so to determine the presence of such influences as practice, fatigue, etc. It may also be applied to a table of correlations to determine whether among the abilities or attributes measured there is any general factor at all. For suppose a number of tests of emotionality be given, then though they may intercorrelate to a certain extent, nevertheless, perhaps owing to the fact that the amount of emotion aroused is specific to the particular presentation there may be no general factor that might rightly be called emotionality.

---

[1] Spearman, C.: "Abilities of Man." 1927, Appendix.

The Application of the Criterion.—For purposes of illustration Table I may be subjected to test by this formula. Using the Tetrad Differences Criterion (as equation (3) is termed) the differences are calculated for all possible tetrads (210 in all) in the table given, *e.g.:*

$$r_{12}r_{34} - r_{23}r_{14} = -.048$$
$$r_{12}r_{35} - r_{23}r_{15} = -.002$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$
$$r_{57}r_{68} - r_{58}r_{67} = \quad .013$$

Their distribution may then be plotted, together with a normal curve, as shown in Fig. 1.[1] If the distribution satisfies this requirement, as in the figure it may be assumed that they are chance deviations from zero, that is that the equation holds, and that *g* and specific factors only are present.[2]

It may be mentioned that tables showing eight measurements (or more) are suited to this treatment. With a lesser number of coefficients the number of differences is too small to give a distribution. Small tables are best treated by showing the individual differences and their probable errors.

Part III. Group Factors and Special Abilities

The Presence of Group Factors.—Having demonstrated the method of application of the Tetrad Differences criterion, it is necessary to

---

[1] Professor Spearman (in a private note) explains the use of the criterion thus: "Suppose a table of correlations to be set out in the ordinary way; each tetrad will form the angles of a rectangle. Then the tetrad difference can quite equally well be calculated in two ways. (*A*) It can be taken as the product of the top left and bottom right values minus the product of the top right and bottom left values. Or (*B*) it can be taken as the product of the top right and bottom left values minus the product of the top left and bottom right ones. In any ordinary table of correlations the fact of being right or left, top or bottom is merely an accident of the way in which the table is set out. Hence a complete frequency distribution must include both, and the distribution must be quite symmetrical. The criterion as to *g* does not consist in seeing whether the distribution is 'normal,' but in seeing whether this observed distribution does or does not show a greater *scatter* than is possessed by a normal distribution having the same area, but having its scatter determined by the theoretical probable error of the tetrad difference. Thomson has distorted my frequency distributions by recalculating them and using only (*A*)." (Thomson: *British Journal of Psychology*, Vol. XVII, p. 235.)

[2] For a series showing other factors see *British Journal of Psychology*, 1926, p. 100.

consider the use which has so far been made of it, and its value for further experimental work. It has been used in England by Spearman (and his associates and students) and by Burt, mainly for the purpose of demonstrating and proving the existence of some General Ability, or general intelligence, as a unitary function, or group of functions acting as a unity. For this purpose it has been necessary to insist that tests which purport to measure this general ability should be tests in which the two factors only are present. This should not result in the conclusion that there cannot be more than two factors or types of factors, for as will be presently shown, there is another type, group factors, the study of which, in a scientific manner, has been made possible by the clarification of concepts resulting from the statistical work described.

According to the argument developed above, abilities in tests such as Analogies, Opposites, etc., may be analyzed. If such subtests in a test be given, and repeated, the analysis would be thus:

$$\text{1st Opposites test score} = g + s_0 + e_1 + e$$
$$\text{2nd Opposites test score} = g + s_0 + e_2 + e$$
$$\text{1st Analogies test score} = g + s_a + e_1 + e$$
$$\text{2nd Analogies test score} = g + s_a + e_2 + e$$

where s_0 and s_a are factors peculiar to Opposites and Analogies respectively, and e_1 and e_2 are factors peculiar to the 1st and 2nd occasions of testing respectively.

Now if the scores in such tests are intercorrelated, a table may be constructed thus:

TABLE II.—INTERCORRELATION OF OPPOSITES AND ANALOGIES, REPEATED ONCE

	O_1	O_2	A_1	A_2
O_1	...	.85	.70	.55
O_2	.85	...	.50	.65
A_1	.70	.50	...	.90
A_2	.55	.65	.90	

The coefficients given are hypothetical. The TD criterion may be applied, with the following results:

$$(X)\ r_{O_1O_2} \cdot r_{A_1A_2} - r_{O_1A_2} \cdot r_{O_2A_1} = .485 \text{ PE} \pm .0027$$
$$(Y)\ r_{O_1O_2} \cdot r_{A_1A_2} - r_{O_2A_2} \cdot r_{O_1A_1} = .305 \text{ PE} \pm .0022$$
$$(Z)\ r_{O_1A1} \cdot r_{O_2A_2} - r_{O_1A_2} \cdot r_{O_2A_1} = .180 \text{ PE} \pm .0033$$

These large differences prove that there are more than two factors present. If the figures are examined, it appears that the difference (X) is due to the fact that $r_{O_1O_2}$ and $r_{A_1A_2}$ are unduly high; that is, the two Opposites tests have something in common peculiar to Opposites; and a similar peculiarity occurs in the case of Analogies. From the analysis above, it is easy to see that it is the factor s_0 which is common to the two Opposites tests, and s_a which is common to the Analogies.

The difference (Z) is due to the fact that $r_{O_1A_1}$ and $r_{O_2A_2}$ are unduly high. Again there must be some factor causing this. The analysis shows that e_1, the factor peculiar to the first occasion of testing is common to both Analogies and Opposites taken on the same first occasion; similarly e_2 for the second occasion.

Causing the difference (Y) are the two above factors acting in opposite directions, the former having the greater weight.

These group factors, which are common to some tests are distinguished from (*a*) the general factor which is common to all, and (*b*) the specific factors which are peculiar to each. If the desire is to obtain a measure of g uninfluenced by factors of any sort, it is necessary to eliminate these group factors. The factors s_0 common to all opposites tests, s_a common to all Analogies tests, s_x common to each test of any type may be eliminated by constructing a test of many forms or tests. Similarly e_1, e_2 . . . e_n may be eliminated by testing on more than one occasion. Over all and common to all in the above test is the verbal factor which has not been eliminated.

From the above demonstration can be seen the reason why Spearman has insisted on the elimination of all group factors, in order to prove the presence of g. For if g were altogether absent there would still be correlation between each of the tests in Table II, owing to the group factors. It can be shown[1] however, (*a*) that the correlations could not be nearly so high, and (*b*) that there is no reason why there should be any relation between the degrees of correlation, if g is altogether absent.

The Elimination of Group Factors.—It is very difficult to eliminate altogether group factors in tests by the method of increasing the number and length of subtests. An alternative method is to admit

---

[1] If g be absent, and the e's uncorrelated then by Thomson's formula,

$$r = \frac{\text{Number of factors common}}{\text{Geometrical mean of total factors}}$$

the average of the coefficients in this table should not be more than .33.

such as cannot easily be eliminated, but to minimize their effect. It will have been noted that a group factor is a commonality among specific factors. If therefore the number of specific factors be reduced to a minimum the effect of the group factors will *ipso facto* be reduced. The figures given in Table II have deliberately shown group factors very clearly. As a contrast, and as showing the reduction of effect by reducing specific factors, by the use of tests well selected according to the principles, the Table III below shows some actual coefficients obtained.[1]

TABLE III.—INTERCORRELATION OF OPPOSITES AND ANALOGIES, REPEATED ONCE
(From Table II, Slocombe: *British Journal of Psychology*, 1926, p. 104)

	O_1	O_2	A_1	A_2
O_1	...	.75	.74	.75
O_2	.75	...	.76	.79
A_1	.74	.76	...	.74
A_2	.75	.79	.74	

The Tetrad Differences obtained from this table are:

$$(X)\ r_{O_1O_2} \cdot r_{A_1A_2} - r_{O_1A_2} \cdot r_{O_2A_1} = -.015 \text{ PE} \pm .008$$
$$(Y)\ r_{O_1O_2} \cdot r_{A_1A_2} - r_{O_2A_1} \cdot r_{O_1A_1} = -.029 \text{ PE} \pm .008$$
$$(Z)\ r_{O_1A_1} \cdot r_{O_2A_2} - r_{O_1A_2} \cdot r_{O_2A_1} = -.014 \text{ PE} \pm .008$$

These differences are all less than the PE, and show the absence of group factors. Using the method of inference, however, logical necessity requires the admission that they are present, but masked by accidental errors. It may be said that the influence of group factors has been reduced to negligible proportion in comparison with the other errors of measurement.[2]

The same argument may be applied to the broader consideration of Linguistic tests and the group factor of verbality. Two sets of results are interesting in this connection.

1. Miss F. Gaw[3] intercorrelated performance tests and linguistic tests, (Binet) and found that "the tests (performance) which correlate

[1] C. S. Slocombe: The Constancy of *g*. *British Journal of Psychology*, 1926, p. 104.

[2] This answers the neglected question raised by T. L. Kelley in "Statistical Method," 1923, p. 202. The results here shown need not necessarily be of general application, for Miss Davey did find that certain tests did contain a group factor of significant size.

[3] *British Journal Psychology*, Vol. XV, p. 374.

best with general intelligence—as measured by the Binet scale—also correlate best with central performance capacity. This seems to furnish strong corroborative evidence that the central factor measured by performance tests is largely general intelligence." Her results unfortunately cannot be put in such a form as to make the criterion applicable.

2. Miss Davey[1] using linguistic tests and a pictorial form of the same material also found no significant group factors peculiar to either. She calculated all TD's from the table of intercorrelations, and found that with the exception discussed in the footnote on p. 18, the distribution of differences was normal, and that there were none so large as to indicate group factors peculiar to either form.

Her results may be summarized thus:

	Verbal	Pictorial
Verbal	.40	.32
Pictorial	.32	.31

The Tetrad Difference is .0218 (*PE* .017).

The present results then justify the use of linguistic tests in the measurement of general intelligence of the majority of groups of individuals; and indicate that even though verbal intelligence may differ from general intelligence, the accidental errors in measurement are at present greater than this difference.

The whole argument has been the necessity of carefully eliminating all factors other than g. General intelligence is then seen to be that which is left, causing correlation between different abilities, when all else has been eliminated. But as in every *single* ability there are two factors, or types of factors, and two only, it is obviously impossible to determine by a knowledge of the amount of the general ability what the specific ability would be. In ordinary algebraic symbols, the value of y in the equation, $y = fx + a$ cannot be determined from a knowledge of x only, where a is any integer, positive or negative, independent of x.

The Measurement of Special Abilities.—The consideration of the results of Miss Gaw and Miss Davey, and of group factors generally indicates the meaning of such terms as social intelligence, moral intelligence, educational ability, and the like. For taking the first named as an example, a set of tests may be devised, involving the educing of

[1] *British Journal Psychology*, Vol. XVII, p. 36.

relations and correlates of what may be termed a social nature, with social material. Then the tests should be found to measure

$$\begin{aligned} \text{1st test} &= g + s_s + s_1 \\ \text{2nd test} &= g + s_s + s_2 \\ &\cdots\cdots \\ &\cdots\cdots \\ N\text{th test} &= g + s_s + s_n \end{aligned}$$

$s_1, s_2 \ldots s_n$, being specific factors are eliminated on addition. s_s, the specific factor of sociality being included in all tests, thereby being a group factor, is additive, and so the total score will be $(g + s_s + e)$ where e is such error as may be not eliminated. In order to obtain the social ability it will be necessary to give also a test of general intelligence, giving a score of $(g + e)$, and adjusting to suitable units, in terms of SD subtract the latter, when the score in sociality remains, affected only by some error in measurement.

It may be asked why it is necessary to eliminate g. Is it not sufficient if the ability to apply intelligence to social situations is measured? There are two reasons why it is better to carry the process to its logical conclusion, one practical, and the other more theoretical. Practically, the majority of tests of various types of intelligence are in a linguistic form, or in some form such that the amount of general intelligence involved is very high. Consequently the element of sociality is rather overshadowed by the general ability, and small differences in the latter mask differences in the former, so that the test fails to differentiate on the basis of sociality. Theoretically, it would be interesting to know whether there is any social "thing" which might be so common as to become a group factor. That is to say, may it not be that sociality as displayed is specific to the occasion, circumstance or stimulus, and is not possessed of any generality. By the means considered above, it is possible to determine whether such a "thing," function or faculty, can be correctly regarded as existing. If g be partialled out, then if there remain correlation between the test scores, it is shown that there is some such fairly general function. Incidentally, it may be mentioned that such tests of sociality and the like, are often criticized because they correlate with intelligence. But it should be clear that it is almost impossible to get any test which does not correlate with intelligence to some extent. (That is why the adjective general is applied to the word Intelligence.) This fact does not prove that the test does not measure

sociality, morality, or what not, but merely that the measurement is influenced by the large amount of g involved.

One further point in connection with such tests of group factors, shows clearly the value of the method discussed and indicates how all tests of special abilities may be improved in their use. Suppose a test to be given of social intelligence, and the subjects be graded by the scores as A, B, C, D, E. Let them also be tested with a general intelligence test and graded 1, 2, 3, 4, 5. Then assuming that g and s_s have equal weight in determining the scores, the possibilities involved in the grades A, B, C, D, E, may be set out thus:

Subject number	Social test grade	May be due to	
		General intelligence (g)	Element of sociality (s_s)
1	*A*	1	*a*
2	*B*	1	*c*
3	*B*	2	*b*
4	*B*	3	*a*
5	*C*	1	*e*
6	*C*	2	*d*
7	*C*	3	*c*
8	*C*	4	*b*
9	*C*	5	*a*
10	*D*	3	*e*
11	*D*	4	*d*
12	*D*	5	*c*
13	*E*	5	*e*

It will be seen that the grade in the factor or element of sociality does not correlate very highly with the Social test grade. An examination of such analyzed results will permit an interpretation of results, which should be useful in understanding individual subjects. To take examples: No. 5 is actually as social as he can be, even though the Social test results do not reveal the fact. No. 2 on the other hand, though he scores highly in the test is not really more than average socially. No experimental results have been worked out in this way, though a somewhat similar logical argument is frequently used in explaining a way inaccuracies in test results, in the diagnosis of abilities or disabilities. The method appears to be of very wide application to all special tests; social, moral, scholastic, musical, vocational, trade, etc.

General Considerations

1. The definition of intelligence here presented is clear and precise. It contains some implications for measurement, which are recognized but not discussed. It attempts definitely to differentiate between those processes, reactions, or responses, in a "total situation" which are of a particular type, from the conditions which enter into and modify their manifestation. The fact that in all abilities intelligence enters as a limiting factor, while implied very generally, is not clearly stated. For example, Hollingworth[1] in discussing correlation and compensation in abilities, does not indicate that intelligence entering as a limiting factor causes some degree of correlation between all abilities, but that the lowness of the correlation and other considerations must incline one to the belief that there is compensation in special abilities. A more general acceptance of the theory here described would lead to a clarification of the understanding of special abilities, and of their compensatory nature.

2. Distinct from special abilities are what may be called attributes of intelligence. The application of the general factor theory (it is, of course, applicable more widely than to intelligence) has revealed two of these attributes, which though independent of the degree of intelligence are as wide in their functioning: (*a*) Perseveration.[2] This is almost synonymous with "lag," a tendency to delay in working up to full use of intelligence, and for the process to continue in a particular field after the necessity has ceased. (*b*) Oscillation.[3] This is a tendency for the variation in intellectual output of an individual to be independent of the nature of the activity.

These factors will be further elucidated when more work has been done in the investigation of the quantitative principles. They are mentioned here, in order to indicate the applicability of the general factor theory, and to illustrate the necessity for a clearer understanding of intelligence, as distinct from perseveration, oscillation, speed, etc.

3. The analysis of test items in terms of relation and correlate eduction, (as will have been concluded from the examples given above)

---

[1] Hollingworth, H. L.: "Mental Growth and Decline." 1926, p. 30.

[2] Lankes, W.: Perseveration. *British Journal Psychology*, Vol. VII, p. 387.
Bernstein, E.: Quickness and Intelligence. *British Journal Psychology*, Monograph Supplement No. 7.

[3] "Abilities of Man." *Loc. cit.*, p. 322.

is a laborious proceeding.[1] Nevertheless, it is essential to the building of a test, in which these processes are crucial, and exceedingly valuable in clarifying the concepts, "levels," "altitude" and "width" of intellect.

A short, very interesting set of questions capable of illustrative analysis is that given by Courtis.[2] If his first three questions are analyzed in this manner, there appears a definite relationship between the percentage of children failing, and the number of relations or correlates to be educed. His last example, where the number and complexity of the processes is comparatively small, exemplifies the influence of fundaments in determining successful eduction.

4. The psychological analysis (in Part I) of cognitive processes under these principles may be regarded as ultimate, and as deriving considerable advantage in application from that fact. It may be contrasted with analyses which result in definitions in terms of "abstract thinking," "generalization," "analysis and synthesis," "reasoning" and so on. These terms do not describe processes which are distinct, and cannot on account of that fact assist in the selection of tests to be included in a measure of intelligence. For, does not the process of selection involve analysis? Does not analysis involve abstraction? Is not generalization impossible without analysis and abstraction? And do not all these processes, (if they can be called such) involve reasoning?

Spearman has shown that all these processes can best be reduced to the terms of his Principles, and that such reduction clearly indicates their similarities, differences, and the extent to which they coincide.

5. Throughout the whole of this paper, a deliberate and careful avoidance of all physiological terms and analogies has been maintained. This is not because a physiological basis, or the existence of a human entity, which functions as a unity, is not admitted, but because of the difficulty of obtaining satisfactory physiological descriptions. Spearman, in spite of his statement[3] that, "the qualitative system appears to offer no grip at all for explanations of physiological order," nevertheless uses physiological concepts, speaking of general intelligence, *g*, as some sort of "mental energy," which flows through neural paths.

---

[1] Strasheim, J. J.: "A New Method of Mental Testing." 1926.

[2] Courtis, S. A.: Construction of Measuring Instruments in the Field of Education. *Scientific Monthly*, 1925, p. 280.

[3] "The Nature of Intelligence and Principles of Cognition." P. 345.

Thorndike[1] does also, and regards the central factor, which he calls Intellect, *C*, as a capacity to form connections. It really seems however in the present state of knowledge that physiological descriptions are highly speculative, and do not assist greatly in clarification. They certainly form the basis of the major portion of controversy, and for that reason are to be avoided if psychologists desire to cease wrangling about opinions."

In Part I the theory of general intelligence as consisting in the ability to apprehend experiences, to educe relations between fundaments, and to educe correlative characters from fundaments and relations, is explained and illustrated.

In Part II the existence of this ability, the difficulties of its measurement because of the factors which enter into and modify its manifestation, together with the means taken to eliminate these factors, are considered. The statistical procedures by which the successful accomplishment of this elimination is tested, are demonstrated by their application to results.

In Part III special abilities are considered as types of ability more limited in their range. Their influence in tests of general ability, and the impossibility of their complete elimination is demonstrated. Finally are shown methods by which they may be isolated and measured.

---

[1] "The Measurement of Intelligence." 1926, p. 422.

THE RELATIVE INFLUENCE OF FATIGUE ON PUPILS WITH DIFFERENT LEVELS OF MENTALITY

GRAYSON N. KEFAUVER

University of Minnesota

The Problem

The study reported herein was suggested by the teachers of the special classes for the slow moving pupils in the schools of Fresno, California. They believed that pupils with low mentality tire more quickly than normal pupils and that their ability to do class work is lowered because of this fatigue. This would indicate that pupils of low mentality are also low in endurance and in their ability to concentrate, to the extent that a small amount of fatigue is injurious to their work. In this discussion, the term fatigue refers to that factor which causes the reduction in speed and efficiency of work when an individual applies himself to a task for an extended period of time. Does the work of the dull pupil suffer more than that of the bright pupil because of fatigue? Also, what is the relative effectiveness of a stimulus and a rest period in removing the influence of fatigue for the different ability groups?

Organization of the Study

Ninety-eight pupils with intelligence quotients ranging from 60 to 129 were selected from four classes. The normal and bright pupils were in Grade V and those of lower mentality were in two of the special, ungraded classes. They were at 11, 12, and 13 years of age. To facilitate consideration, they have been divided into four groups, using the intelligence quotient as the basis for grouping. The group with IQ 60–79 had 15 pupils; IQ 80–99, 38 pupils; IQ 100–114, 34 pupils; and IQ 115–129, 11 pupils.

Eight of the ten forms of the Los Angeles Mixed Fundamentals Test in Arithmetic were given to the groups at one sitting. The ten forms of this test have the same type of operations and are equal in difficulty. The time for each test was reduced from 20 to 15 minutes so as to make speed an element in the results of the tests. This made a total of two hours of continuous work for the series of eight tests. At the close of the time allowed for each test, the pupils were immediately instructed to turn over the sheet and start on the next test.

They were also directed not to turn to the next test until they were told to do so. At the beginning of the eighth test, the pupils were told that they would have a recess at the close of the test. This factor is introduced to show the extent this stimulus offset the influence of the fatigue resulting from performing the previous seven tests. The pupils returned to their classrooms after 20 minutes of play and took Forms 9 and 10. These two tests were given after the period of play to measure the extent to which the different groups were able to regain their normal efficiency during the period of rest and recreation. All of the tests were given by the regular classroom teacher to insure normal class conditions. Also, the tests were given at the opening of school in the morning so as to be certain that the pupils were not already fatigued at the beginning of the testing period.

Findings

(*A*) *The Influence on Speed of Work.*—The influence on speed of work, that is, the number of problems attempted, is shown in Fig. 1 by giving the average number of problems attempted by each group of pupils in each of the ten tests. There is a striking similarity in the modification of speed of work for the different groups as they advanced through the series of tests. All of the groups obtained a slightly higher average in the second test than in the first. This increase is followed by a decrease for all of the groups except the one with IQ 115–129, which group did not experience a decrease until the sixth test. The two lowest groups, IQ 60–79 and IQ 80–99 had a gradual and constant decrease after the second test. The group with IQ 100–114 had the greatest irregularity in the number of problems attempted with a definite increase in tests 5 and 7 over that of the preceding tests.

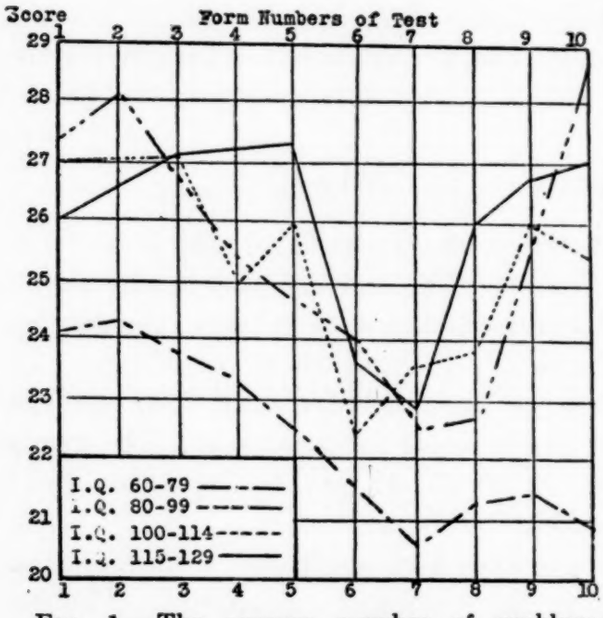


Fig. 1.—The average number of problems attempted by the different ability groups in each of the ten forms of the Los Angeles Mixed Fundamentals Test.

A comparison of the number of problems attempted in the first and seventh tests shows that the group with IQ 80–99 had the largest drop of five points. The slowest group, IQ 60–79 and the group next to the brightest, IQ 100–114 each had a drop of 3.5 points. There is considerable variation among the groups in the number of problems attempted in the different tests, but the retardation of rate of work is not greater for the dull pupils than it is for the bright pupils.

All of the groups responded to the stimulus of being told about the play period after the eighth test. The bright group responded most, bringing the average score up to that of the first test. Also, all of the groups experienced a marked increase in rate after the period of recreation and rest. The slowest group, IQ 60–79, regained least. Only one of the groups, IQ 115–129, worked more rapidly on the first test after the rest period than they did on the first test of the series.

(*B*) *The Influence on the Number of Problems Worked Correctly.*—The influence of fatigue upon the number of problems worked correctly, that is, the score, is shown in Fig. 2. The largest decrease in score was experienced by the brightest group, IQ 115–129. This group had the largest increase on the second test, the largest total drop during the series of tests, and the largest response to being told about the recess following the eighth test, when compared with the variations of the other groups. The slowest group, IQ 60–97, presents the smallest variation in scores. This group made the lowest score of the whole series of tests on the first test, and the variation of the average scores made in the different tests does not exceed one point. The regular shift up and down suggests a somewhat regular alternation of work and relaxation. The reduction of scores from the standard of the first test is very small for groups IQ 115–129 and IQ 80–99, there being only one

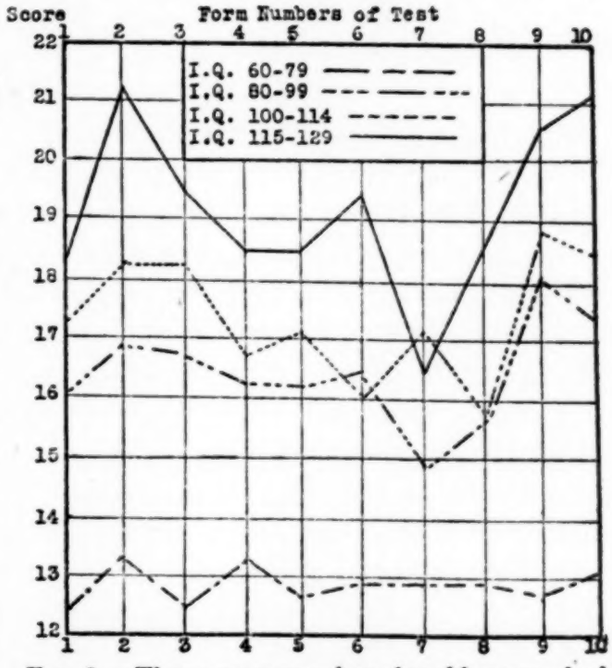


FIG. 2.—The average number of problems worked correctly by the different ability groups in each of the ten forms of the Los Angeles Mixed Fundamentals Test.

test (test 7) for the former group and two tests (tests 7 and 8) for the latter with average scores below that of the first test. However, the reduction of scores is greater for the two brighter groups than for the two groups below average in intelligence. These data indicate that the brighter pupils, rather than the duller, are affected the more by fatigue.

(*C*) *The Influence on Speed and Score Compared.*—There is marked agreement in the averages of the different groups in regard to the influence of fatigue on both speed and score. In general, a reduction in speed of work is associated with a reduction in the average score, but for every group, the reduction in speed is the larger. There is a significant contrast in the relationship between the influence of fatigue on speed and scores for the two extreme groups, IQ 60–79 and IQ 115–129. The brighter group increased in speed in the second and third tests and maintained the high speed during the fourth and fifth tests. In contrast with these variations, there was a definite decrease in score in the third and fourth tests. Also, when the speed of work experienced a very large drop in the sixth test, there was a marked increase in the average score over that of the preceding test. In the slower group, there was a gradual decrease in the speed of work following the second test, but the scores remained practically the same from test to test.

(*D*) *The Variation of Influence on Individuals.*—The treatment of these data by averages for the different groups fails to show the extent of variation of influence among individuals. As would naturally be expected, there is a larger contrast between individuals than between the averages of groups. Do some individuals start to work at their maximum speed and capacity while others increase as they advance through the series of tests? The work of each pupil was recorded graphically so that individual variations could be readily noted. The results were analyzed and four pupils were selected to show the different reactions in speed and score. The records presented in Figs. 3 and 4 indicate that some pupils experience a greater decrease in efficiency than others. Pupil *A* in Fig. 3 had a very large reduction in speed following the second test while pupil *D* continued to increase throughout the series with the exception of the two tests 5 and 8 when he approached the record of the first test. Pupil *C* increased in speed in the second and third tests and did not drop below the first test in any of the subsequent tests. Pupil *N* in Fig. 4 maintained the same score in the first four tests with minor fluctuations in the last six tests. Four of these six tests were above and two below the

standard set in the first tests. Pupils *O* and *P* experienced a marked decrease following the first test and did not make as high a score on any of the other tests of the series as was made on the first test. The contrast of these two pupils with pupil *M* is striking where there is a marked increase in the second test and in only one test, test 7, does the score fall below that of the first test.

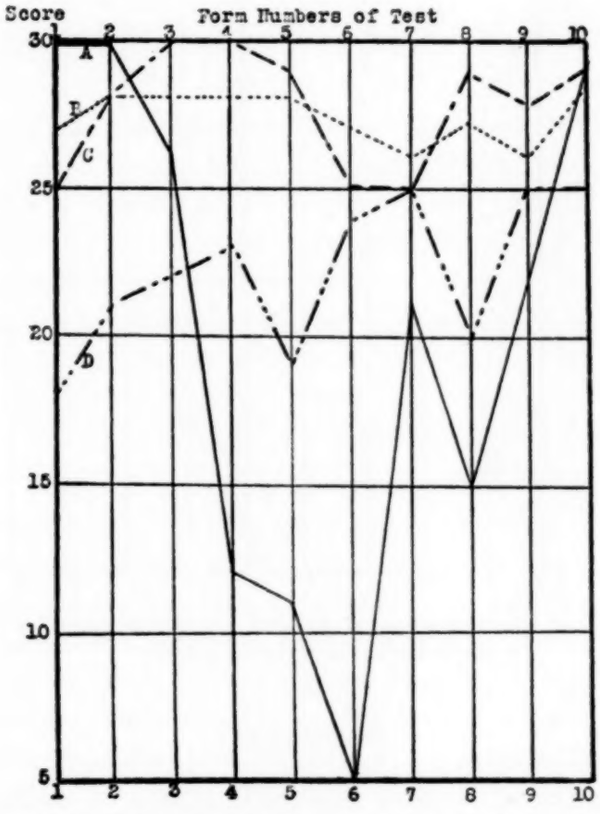


FIG. 3.—Variation in the influence of fatigue on the number of problems attempted by individual pupils.

These data furnish some support to the belief that certain pupils begin to work with a low standard of efficiency and increase after working for a period of time; that others begin with a maximum of drive and efficiency but soon experiences a reduction in efficiency; and that others, in contrast with these two extremes, begin work at a certain degree of efficiency and maintain it over an extended period of time. It cannot be said that pupils distribute themselves into types as they are generally distributed between the two extremes. It was not possible to obtain a type of record which was typical of any of the intelli-

gence-quotient groups as large variations were in evidence in each of the groups.

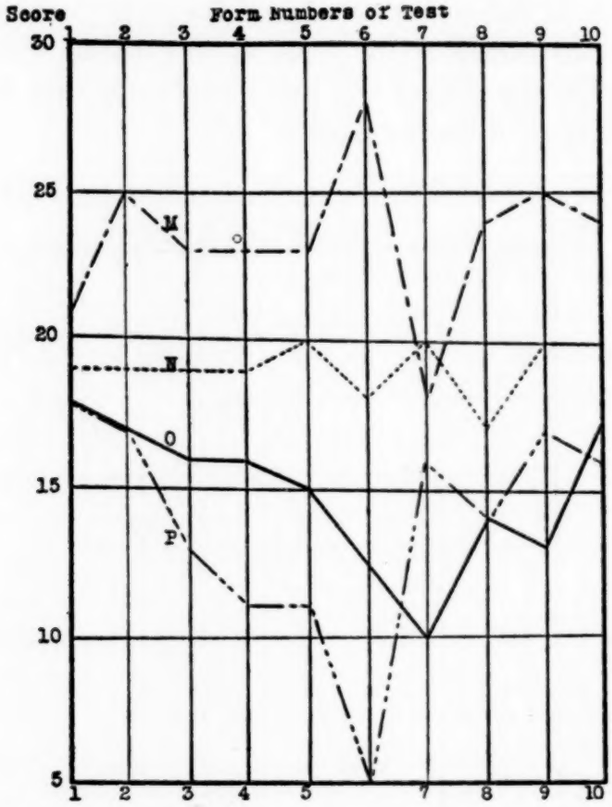


FIG. 4.—Variation of the influence of fatigue on the number of problems worked correctly by individual pupils.

SUMMARY

1. The influence of fatigue upon the speed of work was approximately the same for the brighter as for the duller pupils.

2. The influence of fatigue upon the scores was greater for the brighter than for the duller pupils.

3. The influence of fatigue can be partly removed by presenting a challenging stimulus.

4. A twenty-minute period of rest and recreation was sufficient for the pupils to throw off much of the fatigue resulting from the two hours of continuous work.

5. The influence of fatigue was larger on the speed of work than on the average scores obtained in the different tests.

A STUDY OF CAUSES OF RETARDATION AMONG MEXICAN CHILDREN IN A SMALL PUBLIC SCHOOL SYSTEM IN ARIZONA[1]

O. K. GARRETSON

University of Texas

The purpose of this study is to determine, by objective methods, some of the factors which lead to retardation among Mexican public school children. The total number of pupils in the school system involved in the investigation was 421. Due to various causes, complete data could be procured from only 314 of this number. For this reason the results in the pages which follow are based upon the latter number of children. Of this number 197 were American and 117 were Mexican pupils.

The American or white group comprises approximately 60 per cent of the school enrollment and consists principally of children of northwestern-European ancestry whose parents are native to the United States. There is, however, a very small percentage of Italians and Austrians. English is the language spoken in the homes of the children of Austrian and Italian parentage and they are accepted as equals by the remainder of the group.

The principal industry of the community is the smelting of copper and all of the families are directly or indirectly dependent upon this industry for their income. The highest monthly wage of any of the parents of children composing the white group is $250 and the mean wage, between $125 and $150. None of the parents has a university degree, and not more than 25 per cent have completed a four-year high school course.

The Mexican group furnishes about 40 per cent of the total enrollment in the school system. The mean income of the Mexican parent is, when employed, about $110 per month. Probably not more than 80 per cent of the heads of the Mexican families have steady employment. With few exceptions, Spanish is the language spoken in the homes and on the streets of the Mexican part of the town. It is

[1] The complete data upon which this article is based are to be found in a Master's thesis in the Library of the University of Texas, which was worked out by the author under the direction of Dr. C. T. Gray. The author is indebted to Dr. C. T. Gray and Dr. Helen Koch for suggestions upon the formulation of this article.

TABLE I.—SHOWING AGE-GRADE DISTRIBUTION OF AMERICAN PUPILS

Chronological age in months	Grades								Total
	I	II	III	IV	V	VI	VII	VIII	
72– 83	16	..	..	..	..	..	..	..	16*
84– 95	13	10	..	..	..	..	..	..	23
96–107	2	9	15	1	..	..	..	..	27
108–119	1	3	8	8	1	1	..	..	22
120–131	..	3	2	9	12	3	..	..	29
132–143	..	..	1	5	12	10	..	..	28
144–155	..	..	1	0	8	2	12	2	25
156–167	..	..	..	1	1	4	6	3	15
168–179	..	..	..	..	1	0	2	4	7
180–191	..	..	..	..	..	..	1	4	5
192–203	..	..	..	..	..	..	..	0	0
Total.....	32	25	27	24	35	20	21	13	197

* The size of these groups is due to tests having been given in February to children admitted to the first grade in September with a required chronological age of seventy-two months.

TABLE II.—SHOWING AGE-GRADE DISTRIBUTION OF MEXICAN PUPILS

Chronological age in months	Grades								Total
	I	II	III	IV	V	VI	VII	VIII	
72– 83	7	..	..	..	..	..	..	..	7*
84– 95	12	2	..	..	..	..	..	..	14
96–107	1	4	6	1	..	..	..	..	12
108–119	3	6	4	1	..	..	..	..	14
120–131	..	5	0	7	4	..	..	..	16
132–143	..	1	2	3	2	..	..	..	8
144–155	..	2	0	1	3	3	1	..	11
156–167	..	1	3	1	3	3	2	..	13
168–179	..	..	1	3	3	1	0	..	8
180–191	..	..	2	0	3	3	0	3	11
192–203	..	..	..	1	0	0	0	2	3
Total.....	24	21	18	18	18	10	3	5	117

* The size of these groups is due to tests having been given in February to children admitted to the first grade in September with a required chronological age of seventy-two months.

unusual for a Mexican child to be able to speak English when he enters the kindergarten or first grade.

In studying the problem at hand, the first step was the completion of the age-grade distribution charts as shown in Tables I and II. It will be noted from the results shown in Table I that a consistent decrease in the size of the age group occurs after the twelfth year. This decrease may be explained, in part, by the fact that a portion of each of these groups has dropped out of school after completing the eighth grade. Leaving school prior to completion of the eighth grade or before attaining a chronological age of 16 years is an unimportant factor due to the rigid enforcement of the compulsory attendance law. The age distribution of the Mexican pupils, Table II, shows no consistent decrease or increase in the distribution by age groups subject to the action of the compulsory attendance laws.

It will also be seen that Table I indicates a not unusual distribution of the American children by grade enrollment but a marked decrease exists in the sixth, seventh, and eighth grades of the Mexican group as seen in Table II. This is possibly due to the attainment by some members of the group of a chronological age that frees them from the operation of the school attendance laws prior to completion of these grades. This distribution shows that, should each Mexican child complete one grade each year, 24 of the 117 Mexican pupils would reach their sixteenth year prior to enrollment in Grade VIII.

It seems clear that the elimination which occurs in the American group probably removes for the most part the less retarded, more capable pupil but in the Mexican group the effect of elimination is to remove the more retarded and less capable student.

The above facts seemed to make it desirable to study further the problem of elimination in each of the schools by means of certain methods already extant. In compiling such data three standards were employed: (1) That of Ayers[1] which fixes the ages six to eight years as indicative of normal progress for Grade I; (2) the standard suggested by Terman[2] which sets seven and one-half years as the upper limit of the range for Grade I; (3) and a standard making seven years as the age above which a child is considered retarded if still in the first grade. An addition of 12 months for each grade fixes the

[1] Ayers, L. P.: "Laggards in Our Schools." New York, Charities Publications Committee, 1909, pp. 236.

[2] Terman, Lewis M.: "Intelligence of School Children." Boston, New York, Houghton Mifflin Company, 1919, pp. 317.

upper limit for normal progress, as defined by any of these standards, in each of the succeeding grades.

From the results shown in Table III we find that, regardless of the method of accounting, the retardation of the Mexican is greater than that of the American group. This is true of each grade with the exception of the seventh which contains an insufficient number of Mexicans to give a satisfactory sample, there being but three Mexicans in a class of 24 pupils. We find the retardation of the Mexican child, considering the school grade as a unit, is from three to eight times as great as that of the American child, depending upon the standard by which the retardation is measured. The more liberal the standard used, the greater the disadvantage to the Mexican pupil, for purposes of comparison. Raising the age considered indicative of retardation adds more Americans than Mexicans to the group accepted as making normal progress. The above is true when the mean retardation in months is determined and when the percentage of children retarded in each of the two groups is found.

TABLE III.—SHOWING MEAN RETARDATION IN MONTHS OF AMERICAN AND MEXICAN PUPILS

Grades	Method I		Method II		Method III	
	White	Mexican	White	Mexican	White	Mexican
One	.62	4.37	1.59	5.90	3.53	9.25
Two	3.56	14.14	5.32	17.85	8.16	24.09
Three	1.88	18.55	3.11	21.22	5.37	24.27
Four	2.50	13.88	4.70	17.16	8.00	21.27
Five	2.34	15.22	4.45	19.55	7.77	24.15
Six	−2.20	9.60	−1.03	13.80	.50	18.90
Seven	1.19	0.00	2.47	.66	4.66	3.00
Eight	.61	9.60	2.92	15.60	7.23	21.00
Group	1.47	12.00	3.08	15.19	5.72	19.56
Per cent of racial group retarded	17.7	48.8	33.5	60.6	49.00	77.00

The percentage of American children making normal progress (Table IV) is about one-half greater than that of the Mexican. No American child is retarded more than 35 months, though 4.2 per cent of the Mexican pupils exceed 48 months retardation.

TABLE IV.—SHOWING PERCENTAGE OF PUPILS EXHIBITING VARIOUS DEGREES OF RETARDATION

AMOUNT OF RETARDATION	AMERICAN	MEXICAN
Normal	78.3	51.1
Retarded 0–11 months	12.1	11.9
Retarded 12–23 months	3.5	13.6
Retarded 24–35 months	2.0	6.8
Retarded 36–47 months	0.0	10.2
Retarded 48 months or more	0.0	4.2
Accelerated	4.0	.8

One factor which is often considered as contributing to retardation is irregular attendance. Data upon this phase of the problem are shown in Table V. The results are based upon the assumption that a child absent a mean of 4 or more days out of 20 for the preceding months of the school term in which the survey was made may justly be termed irregular. The percentage of pupils attending school irregularly among the Mexicans is 1.2 per cent greater than among the American children. This difference is too slight to be of importance as a cause of the greater retardation among Mexican children but it may be one of many minor contributory causes.

TABLE V.—SHOWING PERCENTAGE OF CHILDREN ATTENDING SCHOOL IRREGULARLY

GRADES	AMERICANS	MEXICANS
One	21.8	16.6
Two	12.0	14.2
Three	22.2	33.3
Four	8.3	5.5
Five	22.8	11.1
Six	5.0	30.0
Seven	19.0	0.0
Eight	15.3	40.0
Group	16.7	17.9

The assertion is sometimes made that a portion of the difficulty in the education of the Mexican child is the problem of transientness. It is of interest, therefore, to note that the results compiled in Table VI show the Mexican pupil to have been enrolled on the average in 1.6 school systems and his American class-mate to have been enrolled in 1.2 school systems. This would seem to indicate that for the school under consideration the transientness of the family as an explanation of the greater retardation of the Mexican when compared with the American child is not tenable. It is possible, however, that the transientness of the family may have prevented both American and Mexican children from entering school in places of very temporary

residence. Any greater tendency to remain out of school on the part of the Mexican pupil under these conditions is probably minimized by the activities of the truant officer, due to the fact that the state and county funds are distributed among the schools on the basis of average daily attendance rather than on the basis of the school census.

TABLE VI.—SHOWING TRANSIENTNESS OF AMERICAN AND MEXICAN CHILDREN AS SHOWN BY THE AVERAGE NUMBER OF SCHOOL SYSTEMS IN WHICH THEY HAVE BEEN ENROLLED

GRADES	AMERICAN	MEXICAN
One	1.3	1.1
Two	1.7	1.8
Three	1.6	1.5
Four	2.5	1.5
Five	2.8	2.0
Six	2.5	1.5
Seven	2.9	2.6
Eight	2.7	1.8
Group	2.2	1.6

A further approach to the problem was made by the use of mental ability tests. Two of these tests—the National Intelligence Test, Scale B, and the Pintner-Cunningham Primary Mental Test—involve the use of language, while a third test, the Meyers Pantomime Group Intelligence Test has the use of language reduced to a minimum.

The above tests were administered in order that some conclusion might be derived from a comparison of the results from the two groups of children as to the part which skill in the use and understanding of English plays among the children of Mexican parentage as an element to be considered in the progress of the Mexican pupils in those curricular activities which involve the use and comprehension of the English language.

A comparison of the median mental ages (Table VII) shows almost invariably a lower mental age for the American pupils on the non-verbal than on the verbal tests.

This may indicate that the mental age norms of the tests are not equivalent and that to reach a given mental age greater ability is required on the non-verbal test. If true, this would make more significant the fact that the median mental ages of the Mexican children of Grades III, IV, and V are higher on the non-verbal than on the verbal tests. Although this suggests that the factor producing the difference in accomplishment may be a language handicap opera-

tive on the Mexican pupils, it may be the result of other factors which have not been measured.

TABLE VII.—SHOWING MEDIAN MENTAL AGE IN MONTHS FOR AMERICAN AND MEXICAN PUPILS

Grade	Americans	Mexicans	Americans	Mexicans
	Pintner-Cunningham (Verbal test)		Pantomime (Non-verbal test)	
One	85.13	68.5	*	*
Two	98.00	89.0	88.0	84.0
	National (verbal test)			
Three	106.0	81.0	102.0	86.0
Four	114.6	102.5	115.0	106.5
Five	132.2	127.0	123.5	140.1
Six	139.5	135.0	126.0	107.0
Seven	147.3	146.0	138.3	126.0
Eight	158.0	138.0	139.0	129.0

* The point scores of Grade I were too low to be translated into mental ages.

TABLE VIII.—INTELLIGENCE QUOTIENT MEDIANS BY GRADES

Grades	American	Mexican	American	Mexican
	Pintner-Cunningham		Pantomime	
One	98.5	76.0	*	*
Two	96.0	79.0	82.0	71.0
	National			
Three	81.0	70.0	72.0	74.0
Four	89.5	78.0	95.0	78.0
Five	98.0	80.0	92.0	91.5
Six	107.3	83.0	95.5	69.3
Seven	96.3	97.0	84.5	83.0
Eight	90.5	70.5	76.0	68.5

* The point scores of Grade I were too low to be translated into mental ages.

The median intelligence quotients (Table VIII) are higher for the Americans than for the Mexicans, Grade VII for the National Intelligence Test and Grade II for the Pantomime Group Test excepted.

This apparently agrees with the conclusion of Young[1] that the native capacity, if we assume that intelligence quotients are indicative of native capacity, of the Mexican pupil is less than that of the American child.

A comparison of the median intelligence quotient for the Mexican pupils shows their rating higher on the verbal tests for Grades II, VI, VII, and VIII, while the reverse is true for Grades III and V. The rating is the same for Grade IV. Grade I is not given for reasons previously stated. This may indicate that one type of test is about as satisfactory as the other for the purpose of measuring mental ability.

A correlation of the point scores made by the Mexican and by the American pupils on the National Intelligence Test and on the Pantomime Group Intelligence Test for Grades III to VIII inclusive and on the Pintner-Cunningham Primary Mental Test and the Pantomime Group Intelligence Test for Grades I and II are shown in Table IX.

TABLE IX.—SHOWING INTER-TEST CORRELATIONS

Measure	Subject Group: National pantomime	Subject Group: Pintner-Cunningham pantomime
	American	
r ± PE	.561 ± .038	.749 ± .038
Frequencies	140	57
	Mexican	
r ± PE	.559 ± .054	.442 ± .081
Frequencies	72	45

If we assume the correlation between the performance of the American division of the group in the verbal and non-verbal tests to be the correlation expected from students who speak and hear only English in the home, it may follow that by a comparison with correlations derived from the performance in these tests of children of foreign birth, or parentage, we may discover the presence or absence of a greater degree of language difficulty.

From the results tabulated in Table IX it appears that in the first and second grades the language factor operates to the disadvantage of the Mexican pupil. But the correlation of Grades III to VIII are so

[1] Young, K.: "Mental Differences in Certain Immigrant Groups." Eugene, University of Oregon Publication, July, 1922.

close that we may assume, other factors remaining constant, that the excessive language difficulty is now less important. It is possible, however, that the elimination noted in Grades VI, VII, and VIII of Table II may remove the linguistically handicapped Mexican pupil or other factors, as environment, may in part compensate this handicap.

In view of the fact that the native capacity of the Mexican student has been advanced as an explanation of this retardation it seems that a correlation between the intelligence quotient and the retardation, although not establishing a casual relationship, might indicate to some extent the validity of this assumption. For the purposes of this correlation the scores of pupils of Grades III to VIII were used and since the correlations of the performance of the pupils on the National and Pantomime tests were almost identical for these grades, the quotients derived from the National Intelligence Test were correlated with the acceleration, or decreasing retardation, of the individuals of the American and Mexican sub-groups.

TABLE X.—SHOWING INTELLIGENCE-ACCELERATION CORRELATION OF THE AMERICAN AND THE MEXICAN CHILDREN OF GRADES III TO VIII INCLUSIVE

MEASURE	AMERICAN	MEXICAN
r ± PE	.626 ± .036	.809 ± .033
Frequencies	125	48

These correlations as seen in Table X indicate that the mental ability of the child is closely related to his grade location. It would be unwise, however, to conclude that intelligence, or native capacity, is the solution of the problem. We can but say that intelligence, as measured by the group test, is closely associated with the degree of retardation, particularly of the Mexican pupils, in the grades under consideration.

CONCLUSIONS

1. The Mexican pupil of this system is retarded by the most liberal accounting, 10.53 months more than his American classmate.

2. By the same accounting, there are 30.01 per cent more individuals of the Mexican group retarded than of the American group.

3. The Mexican of this system is more irregular in school attendance than is the American pupil. The difference is so slight that it is of questionable importance as a factor in the causes of retardation.

4. The results seem to indicate that transientness is not an explanation of the greater retardation of the Mexican children of this group.

5. In this study the verbal tests are about as satisfactory as the non-verbal test for the purpose of measuring mental ability in Grades III to VIII inclusive.

6. The factor of language difficulty operates to the disadvantage of the Mexican in Grades I and II, but is in this group apparently of less importance in Grades III to VIII.

7. Probably the principal factor governing retardation of the Mexican child is his mental ability as measured by the group test.

THE MEASUREMENT OF CHILDREN'S CONCEPTS

LAURANCE F. SHAFFER

The Lincoln School of Teachers College[1]

The important rôle of the basic social concepts in contemporary life was demonstrated by Meltzer in his study "Children's Social Concepts; Their Nature and Development."[2] One of the basic problems of Meltzer's dissertation was to determine the meanings of certain social concepts known by the children. The problem, in his words, was "How can we best tap the contents of the mind so that we may best discover what children know about our list of concepts?" It was solved, after considering many methods and trying several, by the Personal Interview test. This he describes as a controlled association test in which the child responded to the question of the examiner, "What does (name of concept) mean to you? Tell me everything that (name of concept) means to you." The child gave his replies orally, the examiner writing everything that was said. This was an individual examination.

In connection with preliminary trials of a learning experiment in the social studies, the writer was confronted with the problem of adapting the Concept Test to group use. Not only was the time required for the Personal Interview test prohibitive, but it was necessary to test a large number of children immediately after the reading of material designed to teach the meaning of a concept. Group tests of the usual variety, as Multiple Choice and True-false were considered undesirable. Children's meanings are so diverse and sometimes so erroneous that although controlled tests can tell to a certain extent whether children possess a correct meaning, they cannot give adequate clues as to the breadth and quality of meaning or indicate misconceptions.

The purpose of this article is to describe a group Free Expression test which was devised, and to show its relation to the Personal Interview test, its reliability of scoring, and its correlations with controlled group tests.

The final form of the test was as follows: The examiner gave the children a directions sheet, and a blank sheet on which they were to write the responses. The direction sheet read:

---

[1] From the Social Science Research Laboratory of The Lincoln School of Teachers College. Studies under the direction of Harold Rugg.

[2] Columbia University Contributions to Education, Teachers College Series, No. 192 (1925).

Concept Test

We want to find out what certain words mean to children of the different grades. I will read you these words one at a time. It will be your job to write out the ideas that come to your mind when you think of the words that I read. *Do not try to write whole sentences or try to write a composition.* Words and short phrases, if they express your meaning, will do just as well.

Here is an example. Suppose I ask you the question, "What does 'Republic' mean to you? What do you understand by the word 'Republic?'" You will write something like this:

Republic—governed by people—free country—people vote—elect president and congress—President is head of government—Congress makes laws.

And perhaps some other things that "Republic" means to you.

Do you see that it is so much easier to explain what "Republic" means in this way than to write a composition about it? In order to keep the word in mind while you are writing, say it to yourself before every meaning that you write. Say "*Republic*—governed by people—*Republic*—free country . . . etc."

Now we will try another example on the blackboard.

The examiner read the direction sheet aloud, the children reading silently. Then the examiner said, "Now let's see if you can tell me what a word means, expressing the meaning in short phrases and words, without using sentences. I will write your answers on the blackboard." The examiner then asked the class for the meaning of another concept, and asked members of the class to suggest responses, encouraging them by saying "Anything else? Does (name of concept) make you think of any other ideas?" The correct suggestions were written on the blackboard in the form desired.

The examiner then directed the class to take the blank sheet, and showed where to write the responses. Then he asked, "What does (concept) mean to you? What do you understand by the word (concept)?" After 45 seconds he said, "Don't be satisfied with just one or two words, What else does (concept) mean to you?" After 90 seconds he said, "Any other ideas? Do you think of anything else when I say '(concept)'?" After two minutes the examiner said, "That is all on that word. Now for the next one." The same procedure was followed for all of the concepts to be tested.

The use of the example placed on the blackboard, and of the urging while the child was writing were found necessary by experiments. When the former part of the procedure was omitted there was a greater tendency to write sentences rather than words and phrases. With the procedure as described above, 90 per cent of the responses were in the desired form, when it was omitted, only 65 per cent of the

children avoided the set "to write a composition." The urging after 45 and 90 seconds of writing resulted in fuller responses and larger scores.

The Free Expression tests were scored similarly to Meltzer's Personal Interview tests. The scale was

8—Very superior response.
7—Superior response.
6—Good general idea.
5—Some general idea.
4—Correct facts.
3—One correct fact.
2—Confused (both right and wrong responses).
1—Very confused, but some hint of meaning.
0—No response. (Equivalent to "I don't know.")
0—Various complete misconceptions.

In scoring the tests, specimens were first selected which matched each point on the scale. The others were then scored by comparison, in a manner similar to the use of a composition scale. The reliability of scoring, determined by scoring 85 tests twice, making a new comparison scale for the second scoring, was .90 $\pm$ .01.

The correspondence of the group Free Expression test with the Personal Interview test was determined by administering the test in both forms to two groups. In the first group of 72 cases, the individual test was given first, and the group test afterward. The correlation was .67 $\pm$ .05. In the second group of 108 cases, the group test was given first, followed by the individual test. The correlation was .50 $\pm$ .05. When these coefficients are corrected for "attenuation" due to the chance errors in scoring, the reliability of scoring being .90, the probable correlations between the group and individual tests with perfect scoring become .74 and .55.

The value of the group Free Expression test is not, however, entirely measured by these correlations. An analysis of the individual responses shows a striking identity between the thoughts expressed on the individual and group forms. This is seen in the following examples, taken at random from sixth grade responses to the concept "an American." The group test was given first, and is indicated by (*G*), the Personal Interview response being indicated by (*I*). The score of each response is given.

1. (*G*) Person born in America—a foreigner when he gets his papers can be an American. (6)

(*I*) Person born in America—when a foreigner comes over here he gets his citizen papers and can become an American. (6)

2. (*G*) One born in America—one who comes from a foreign country to be an American—speaks English. (6)

(*I*) Anyone born in America—in the United States—one who comes over to America to be an American. (6)

3. (*G*) Born in this country—speaks American language—has American habits and thinks of his country. (6)

(*I*) Means it is a person who is born here—may be of our religion—person who is born in this country. (3)

4. (*G*) Born in America—should love your country—love the flag—fight for your country. (5)

(*I*) You were born in America—you always have to stand by your country—love your country, and not turn against it. (5)

5. (*G*) Born in America—of American parents—an American citizen. (4)

(*I*) Born in America—of American parents. (3)

6. (*G*) A good citizen—good country for Army—Junior Red Cross. (3)

(*I*) Means a good citizen. (3)

7. (*G*) Citizen—a man who is born in America. (3)

(*I*) Citizen. (3)

8. (*G*) Person born in America. (3)

(*I*) Means a person born in America. (3)

9. (*G*) An American citizen—free man—country—man—settler. (3)

(*I*) American man—an American citizen—American flag—large buildings of America. (2)

10. (*G*) Person—religion. (0)

(*I*) Name of a person—a religion. (0)

These meanings seem pitifully meagre, for so familiar a concept, but the group test brings out as much as the Personal Interview. It is in this qualitative aspect that the test is of most value.

The correlations between the Free Expression test and controlled group tests were determined with the concept "The Industrial Revolution." The correlation with a 10 question Multiple Choice test was $.56 \pm .04$, and with a 32 element true-false test, $.54 \pm .04$. The correlation between the two objective tests was $.80 \pm .02$. The correlation of the Free Expression test with a group intelligence test was $.23 \pm .05$ (160 cases). This is somewhat lower than the correlation of the usual group achievement test with intelligence scores.

A MODIFICATION OF TOLLEY AND EZEKIEL'S METHOD OF HANDLING MULTIPLE CORRELATION PROBLEMS

HENRY E. GARRETT

Columbia University

In a recent paper,[1] Tolley and Ezekiel have presented an ingenious method of solving multiple correlation problems when (1) the multiple regression equation, (2) the coefficient of multiple correlation, R, and (3) the partial or "net" correlation of each independent variable with the dependent variable are wanted, but not the various intermediate partial r's of higher orders. The method enables one to calculate—with a minimum of arithmetic—the regression coefficients (the b's) from which the regression equation is readily obtained.

In applying Tolley and Ezekiel's method, the present writer has made several rather simple modifications in it which seem, however, better to adapt it to correlation problems which arise in psychology and education. These changes are given in the present paper, together with enough of the procedure developed by the original authors to make the modifications intelligible. Those who wish to work through the mathematical development of the method should go directly to the paper of Tolley and Ezekiel previously mentioned.

The distinctive feature of this method as outlined by its authors consists essentially in considering each complete record (*e.g.*, of an individual on a number of tests) to be an observation equation in the regression form

$$X_1 = a + b_{12.34\ldots n}X_2 + b_{13.24\ldots n}X_3 + b_{14.23\ldots n}X_4 + \cdots$$

in which (to take the illustration from test psychology) X_1 is the "criterion" score, X_2, X_3, X_4, . . . X_n are a series of test scores, and a and the b's are unknowns to be determined. There will, of course, be as many observation equations as there are individual subjects, and the question arises of finding a set of values which will most nearly satisfy all of them. Assuming that the values of a and the b's which will satisfy all of these equations are those which will make the sum of the squares of the residuals a minimum, we can determine the

[1] Tolley, H. R., and Ezekiel, M. J. B.: A Method of Handling Multiple Correlation Problems. *Journal of the American Statistical Association*, Vol. XVIII, December, 1923, pp. 993–1003.

unknowns by building up "normal" equations on the principle of least squares, and solving these simultaneously.[1]

As finally determined by Tolley and Ezekiel (see page 997, equation (5)) after a series of mathematical transformations, these normal equations from which the b's may be obtained are

$$\begin{aligned}
p_{12} &= \sigma_2^2 b_{12.345\ldots n} + p_{23} b_{13.245\ldots n} + p_{24} b_{14.235\ldots n} + \\
&\qquad p_{25} b_{15.234\ldots n} + \cdots \\
p_{13} &= p_{23} b_{12.345\ldots n} + \sigma_3^2 b_{13.245\ldots n} + p_{34} b_{14.235\ldots n} + \\
&\qquad p_{35} b_{15.234\ldots n} + \cdots \\
p_{14} &= p_{24} b_{12.345\ldots n} + p_{34} b_{13.245\ldots n} + \sigma_4^2 b_{14.235\ldots n} + \\
&\qquad p_{45} b_{15.234\ldots n} + \cdots \\
p_{15} &= p_{25} b_{12.345\ldots n} + p_{35} b_{13.245\ldots n} + p_{45} b_{14.235\ldots n} + \\
&\qquad \sigma_5^2 b_{15.234\ldots n} + \cdots \qquad (1)
\end{aligned}$$

in which p_{12} = the mean product sum $\frac{\Sigma(x_1 x_2)}{N}$, $p_{13} = \frac{\Sigma(x_1 x_3)}{N}$ etc., and the σ's are the standard deviations of the variables X_2, X_3, X_4, etc. The mean product sums (the p's) are obtained from the correlation tables. When the values of the p's and the σ's are put into the normal equations, these may be solved simultaneously to give the b's which, in turn, are put into the regression equation for "predicting" X_1 from the scores X_2, X_3, X_4, . . . X_n.

The equations given above in (1) are in the form in which they were left by Tolley and Ezekiel. It is clear, however, that $p_{12} = r_{12}\sigma_1\sigma_2$ $\left(\text{since } r_{12} = \frac{\Sigma(x_1 x_2)}{N\sigma_1\sigma_2}, \text{ and } p_{12} = \frac{\Sigma(x_1 x_2)}{N} \text{ by definition}\right)$, that $p_{13} = r_{13}\sigma_1\sigma_3$, etc., and if these substitutions are made for p in the normal equations (1) they become

$$\begin{aligned}
r_{12}\sigma_1 &= \sigma_2 b_{12.345\ldots n} + r_{23}\sigma_3 b_{13.245\ldots n} + r_{24}\sigma_4 b_{14.235\ldots n} + \\
&\qquad r_{25}\sigma_5 b_{15.234\ldots n} + \cdots \\
r_{13}\sigma_1 &= r_{23}\sigma_2 b_{12.345\ldots} + \sigma_3 b_{13.245\ldots n} + r_{34}\sigma_4 b_{14.235\ldots n} + \\
&\qquad r_{35}\sigma_5 b_{15.234\ldots n} + \cdots \\
r_{14}\sigma_1 &= r_{24}\sigma_2 b_{12.345\ldots n} + r_{34}\sigma_3 b_{13.245\ldots n} + \sigma_4 b_{14.235\ldots n} + \\
&\qquad r_{45}\sigma_5 b_{15.234\ldots n} + \cdots \\
r_{15}\sigma_1 &= r_{25}\sigma_2 b_{12.345\ldots n} + r_{35}\sigma_3 b_{13.245\ldots n} + r_{45}\sigma_4 b_{14.235\ldots n} + \\
&\qquad \sigma_5 b_{15.234\ldots n} + \cdots \qquad (2)
\end{aligned}$$

These equations are in a distinctly more usable form than those given by Tolley and Ezekiel for the following reasons: In the first

[1] A simple treatment of least squares will be found in Gavitt's: "A First Course in Statistical Method." New York, 1925, pp. 190–192.

place in all of the printed correlation charts which make use of the "standard" method, deviations in X and Y are taken from "guessed averages" in units of step-interval, and the p (*i.e.*, the mean product sum) is not directly obtainable; in the second place in several printed correlation charts widely used by psychologists (Otis' and Toops' are examples) the product sums are not found at all, and hence the method is entirely inapplicable here; finally the equations in (1) make it imperative that we have the actual correlation charts, and are, therefore, of little or no value so far as most published correlational data is concerned. In the equations as given in (2) only the intercorrelations and the σ's of zero order are needed in order to solve for the regression coefficients.

The equations in (2) may be expressed in a somewhat simpler form if we divide both sides of each equation by σ_1; put $R_{12} = \frac{\sigma_2}{\sigma_1}$, $R_{13} = \frac{\sigma_3}{\sigma_1}$, $R_{14} = \frac{\sigma_4}{\sigma_1}$, $R_{15} = \frac{\sigma_5}{\sigma_1}$, etc., and let $B_{12} = b_{12.345 \cdots n}$, $B_{13} = b_{14.245 \cdots n}$, $B_{14} = b_{14.235 \cdots n}$, $B_{15} = b_{15.234 \cdots n}$, etc. Making these substitutions we have

$$\begin{aligned}
r_{12} &= R_{12}B_{12} + r_{23}R_{13}B_{13} + r_{24}R_{14}B_{14} + r_{25}R_{15}B_{15} + \cdots \\
r_{13} &= r_{23}R_{12}B_{12} + R_{13}B_{13} + r_{34}R_{14}B_{14} + r_{35}R_{15}B_{15} + \cdots \\
r_{14} &= r_{24}R_{12}B_{12} + r_{34}R_{13}B_{13} + R_{14}B_{14} + r_{45}R_{15}B_{15} + \cdots \\
r_{15} &= r_{25}R_{12}B_{12} + r_{35}R_{13}B_{13} + r_{45}R_{14}B_{14} + R_{15}B_{15} + \cdots
\end{aligned} \tag{3}$$

If now we substitute in equations (3) the values of the r's, we may solve them simultaneously for $R_{12}B_{12}$, $R_{13}B_{13}$, $R_{14}B_{14}$, etc.; and from these the regression coefficients are easily obtained.

Should it be desirable to "predict" X_2 from X_1, X_3, X_4, . . . X_n, the normal equations take the form (see Tolley and Ezekiel, page 998, equation (6))

$$\begin{aligned}
p_{21} &= \sigma_1^2 b_{21.34 \cdots n} + p_{13}b_{23.14 \cdots n} + p_{14}b_{24.13 \cdots n} + \cdots \\
p_{23} &= p_{13}b_{21.34 \cdots n} + \sigma_3^2 b_{23.14 \cdots n} + p_{34}.b_{24.13 \cdots n} + \cdots \\
p_{24} &= p_{14}b_{21.34 \cdots n} + p_{34}b_{23.14 \cdots n} + \sigma_4^2 b_{24.13 \cdots n} + \cdots
\end{aligned} \tag{4}$$

Replacing the p's and b's as shown in (2) and (3) above and putting $R_{21} = \frac{\sigma_1}{\sigma_2}$, $R_{23} = \frac{\sigma_3}{\sigma_2}$, $R_{24} = \frac{\sigma_4}{\sigma_2}$, etc., we have

$$\begin{aligned}
r_{21} &= R_{21}B_{21} + r_{13}R_{23}B_{23} + r_{14}R_{24}B_{24} + \cdots r_{1n}R_{2n}B_{2n} \\
r_{23} &= r_{13}R_{21}B_{21} + R_{23}B_{23} + r_{34}R_{24}B_{24} + \cdots r_{3n}R_{2n}B_{2n} \\
r_{24} &= r_{14}R_{21}B_{21} + r_{34}R_{23}B_{23} + R_{24}B_{24} + \cdots r_{4n}R_{2n}B_{2n}
\end{aligned} \tag{5}$$

The normal equations for predicting X_3 or X_4 may be formed in the manner shown above for X_1 and X_2.

The equation for the coefficient of multiple correlation (R) is given by Tolley and Ezekiel (page 998, equation (8), and page 1000, equation (12)) as follows:

$$R^2_{1(234\ldots n)} = \frac{b_{12.34\ldots n}p_{12} + b_{13.24\ldots n}p_{13} + b_{14.23\ldots n}p_{14} + \cdots}{\sigma_1^2} \quad (6)$$

Substituting $r_{12}\sigma_1\sigma_2$ for p_{12}, $r_{13}\sigma_1\sigma_3$ for p_{13}, etc., cancelling the common term (σ_1) in numerator and denominator, and substituting B_{12} for $b_{12.34\ldots n}$, B_{13} for $b_{13.24\ldots n}$, we have

$$R^2_{1(234\ldots n)} = \frac{B_{12}r_{12}\sigma_2 + B_{13}r_{13}\sigma_3 + B_{14}r_{14}\sigma_4 + \cdots}{\sigma_1} \quad (7)$$

Since the b's have already been found from the equations in (3) and the r's and σ's are given, this equation is readily solved for $R_{1(234\ldots n)}$.

The "net" or partial correlation of each independent variable with the "criterion" X_1 may be found from the equation (Tolley and Ezekiel, p. 998 (7))

$$r_{12.34\ldots n} = \sqrt{(b_{12.34\ldots n})(b_{21.34\ldots n})} \quad (8)$$

in which the regression coefficients are obtained directly from equations (3) and (5).

Tolley and Ezekiel do not give a formula for calculating the error made in "estimating" X_1 from the regression equation, *viz.*, $\sigma_{est.\ x_1}$. However, since $R^2_{1(234\ldots n)} = 1 - \frac{\sigma^2_{1.234\ldots n}}{\sigma_1^2}$, and both R and σ_1 are known, we may calculate $\sigma_{1.234\ldots n}$ (which is the same as $\sigma_{est.\ x_1}$) from the formula:

$$\sigma_{est.\ x_1} = \sqrt{\sigma_1^2\,(1 - R^2_{1(234\ldots n)})} \quad (9)$$

If only the regression equation, the coefficient of multiple correlation, and the standard error of estimate are wanted, a 4-variable problem may be solved, with the aid of a calculating machine, in about 35 minutes. Tolley and Ezekiel state that a 10-variable problem can be done in half a day.

To illustrate the use of the formulas given above, the solution of a 4-variable problem taken from the field of test psychology is appended.

Illustrative Problem

Let:

1 = score on the Thorndike Intelligence Examination for High School Graduates.

2 = score on the Digit-symbol Substitution Test.

3 = score on the Turkish-English Vocabulary Substitution Test.
4 = score on the Code Learning Test.

$M_1 = 85.15$ $\sigma_1 = 13.03$ $r_{12} = .09$ $r_{23} = .29$
$M_2 = 287.92$ $\sigma_2 = 47.10$ $r_{13} = .37$ $r_{24} = .57$
$M_3 = 161.53$ $\sigma_3 = 42.91$ $r_{14} = .31$ $r_{34} = .39$
$M_4 = 196.00$ $\sigma_4 = 41.01$

$$R_{12} = \frac{\sigma_2}{\sigma_1} = 3.615$$

$$R_{13} = \frac{\sigma_3}{\sigma_1} = 3.293$$

$$R_{14} = \frac{\sigma_4}{\sigma_1} = 3.147$$

Substituting the values of the r's in equation (3) we have

$$\begin{aligned} .09 &= R_{12}B_{12} + .29R_{13}B_{13} + .57R_{14}B_{14} \\ .37 &= .29R_{12}B_{12} + R_{13}B_{13} + .39R_{14}B_{14} \\ .31 &= .57R_{12}B_{12} + .39R_{13}B_{13} + R_{14}B_{14} \end{aligned}$$

Solving these three equations simultaneously, we get

$R_{12}B_{12} = -.16$ $B_{12} = -.04$ (*i.e.*, $b_{12.34}$)
$R_{13}B_{13} = .31$ $B_{13} = .09$ (*i.e.*, $b_{13.24}$)
$R_{14}B_{14} = .28$ $B_{14} = .09$ (*i.e.*, $b_{14.23}$)

Substituting the values of the regression coefficients in the regression equation

$$x_1 = b_{12.34}x_2 + b_{13.24}x_3 + b_{14.23}x_4$$

we have

$$x_1 = -.04x_2 + .09x_3 + .09x_4 \text{ (Deviation Form)}$$

or

$$X_1 = -.04X_2 + .09X_3 + .09X_4 + 64.5 \text{ (Score Form)}$$

To find $R_{1(234)}$ substitute the values of the B's, r's, and σ's, in (7):

$$R^2_{1(234)} = \frac{-.04 \times .09 \times 47.10 + .09 \times .37 \times 42.91 + .09 \times .31 \times 41.01}{13.03}$$

$$R_{1(234)} = .43$$

The standard error of estimate, $\sigma_{est.\ x_1}$, is obtained from (9):

$$\sigma_{est.\ x_1} = \sqrt{(13.03)^2(1 - .43)^2}$$
$$= 11.76$$

LEARNING CURVES OF BOYS OF IQ'S 76–148

FRANK T. WILSON

State Normal School, Buffalo

Through the courtesy of the teachers and authorities of the training school of the Southern Branch of the University of California a series of experiments in learning was carried on with 20 boys of ages 8½ to 12¾ years, of IQ's 76 to 148 (by the individual Stanford Revision test given by the school counsellor), in Grades III to VI and the opportunity room for superior children. The star tracing mirror experiment and the Witmer cylinder test were used. In nearly every case the subjects made 20 tracings, took 16 cylinder tests, worked on three successive days, at practically the same hour of the day, were in the same room, and performed the tests in the same order. Except in a very few instances no one was present but the writer who gave all the tests. All the boys but one said that they had not done any mirror drawing or mirror writing of any kind, and that they did not practice or talk about the experiments between test periods. No evidence appeared that their word was in any way unreliable. One boy had drawn a star a year or more earlier he thought, but his results showed little if any effects of such remote practice, although effects may have been present.

The purpose of the experiment was to see if any conspicuous differences or marked characteristics would appear at any or at all IQ levels. It was believed that the two tests demanded abilities quite different and that one or the other or both might show characteristic reactions among distinctly different groups.

In explanation of the two tests themselves the following descriptions of the administration of each under the conditions of this experiment are given.

1. *Star Tracing.*—The Starch six pointed star was used as shown on page 30, "Experiments in Educational Psychology," D. Starch, Macmillan Company, 1911. The boys were seated at a low table, made as comfortable as possible, perfect mimeographed stars were provided, a fresh copy was used at each trial, the sheets were held in place by thumb tacks and the surface below was kept smooth by frequently changing card board and other paper, so that impressions could in no way guide the pencil. On the first day 4 tracings were made, the first one being of only the first half of the star. On the second day there

were six tracings, interrupted between the third and fourth by four cylinder tests. On the third day there were ten, interrupted between the fifth and sixth by six cylinder tests. Emphasis throughout the tracing was laid on keeping the pencil on the paper and keeping on the line. After the first few trials there was little need to caution any subjects on the former condition, but throughout the series constant admonition was given to "Keep right on the line." Cautioning was confined to slight variations of the phrase quoted and all suggestions as "Be careful" or "Go slowly" were rigidly avoided.

The experimenter sat at the left and behind the subject who knew, however, that his work was being observed. Subjects were encouraged, without being told to, to talk and their words—sentences usually expired before completion as attention was so concentrated on the tracing—were taken down in short-hand. After each trial the time was recorded. To offset the emphasis on speed the merit of each tracing was briefly discussed with the subject after its completion, such things as cutting corners and points and other poor tracing being mentioned as well as particularly good places. As above noted there was also insistent cautioning during the trials to keep on the line. In this effort of the experimenter to secure good work care was taken not to upset the subjects emotionally, but so to treat each boy that he would do his best to make good stars. Without elaborating further the writer believes that even in this subjective way the variables in ideals, effect, and emotions were kept relatively uniform for all subjects as far as influencing factors outside the subjects were concerned.

2. The Witmer Cylinder is a circular wooden box in the circumference of which are 18 holes of different diameters or depths or both, and so arranged that in either direction there is a certain sequence. To fit these holes there are cylinders which for the experiment are mixed up in the central space inside the circle of holes. The task is to put the cylinders into the holes where they belong as rapidly as possible without making errors. The subjects were shown how certain cylinders "just fit" certain holes and how they are too long, short, thick or thin for any others. An error was defined as an attempt to put a cylinder where it did not belong. Time and errors were recorded. To offset emphasis on time the number of errors was counted and given the subject after each trial and at the start of the next one encouragement to "do it now without any mistakes" was given. As a matter of fact subjects counted their errors as they went along. Six trials were given the first day, four the second, and six the third.

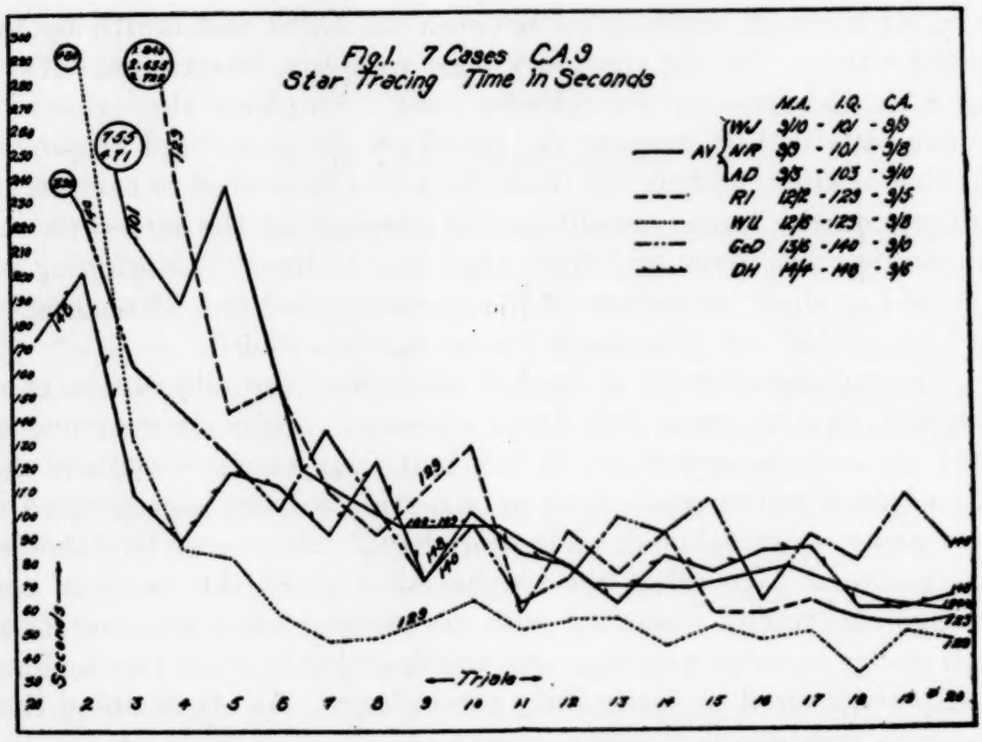


Fig. 1. 7 Cases C.A. 9
Star Tracing Time in Seconds
M.A. I.Q. C.A.
AV WJ 9/0 - 101 - 9/3
AV NR 9/3 - 101 - 9/8
AV AD 9/3 - 103 - 9/10
RI 12/2 - 129 - 9/5
WU 12/8 - 129 - 9/8
GeD 13/5 - 140 - 9/0
DH 14/4 - 148 - 9/6
Seconds
Trials

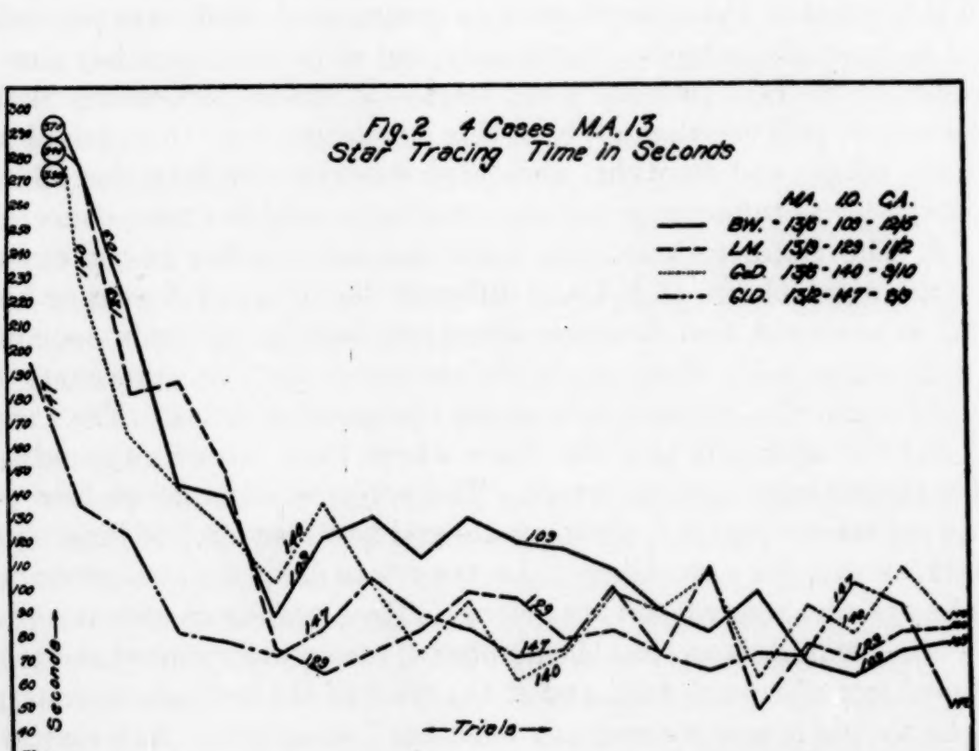


Fig. 2 4 Cases M.A. 13
Star Tracing Time in Seconds
M.A. I.Q. C.A.
BW. 13/8 - 103 - 12/5
LM. 13/9 - 129 - 11/2
GeD. 13/8 - 140 - 9/10
GLD. 13/2 - 147 - 8/9
Seconds
Trials

Records

1. *Star Tracing.*—As it was attempted to keep accuracy or merit constant, time is the variable and results are therefore given in seconds. For the sake of brevity a few curves are given in the accompanying figures which are typical of all the results secured. Figure 1 shows the curves for seven boys ranging in age from 9 years 0 months to 9 years 9 months, in IQ from 101 to 148, and in mental age from 9–3 to 14–4. One line in the figure shows the average for three boys of IQ's 101–103. One characteristic is evident: Time decreases with practice, rapidly at first, then less and less so, with many ups and downs and crossing and recrossing of lines.

Figure 2 gives the curves for four boys ranging from 8 years 9 months to 12 years 5 months in age, from 109 to 147 IQ, and in mental age from 13–2 to 13–9. The figure presents much the same features as Fig. 1.

Figure 3 gives the curves for five boys 11 years 0 months to 11 years 8 months of age, from 90 to 147 IQ and in mental age from 10 years 5 months to 17 years 2 months. In this figure the curve of F. D. (IQ 90) suggests that learning is slower for him after 10 or a dozen trials, despite a very good start—but he comes in with the field at the end, less than 10 seconds behind IQ 147, who is chronologically slightly older!

When Figs. 1, 2, and 3 are superimposed upon one another it is found that except for the beginnings of the curves for the younger boys (which are higher in the majority of these few cases) there are no marked differences. All curves fall rapidly for five or six trials, less and less so thereafter, and go up and down and cross and recross with no apparent consistency.

2. *Cylinder Test.*—A composite score made up of the sum of the number of seconds plus five times the number of errors was used as a means of comparing results. Figure 4 shows the curves according to this composite score for the same boys as Fig. 1. Two of these subjects had but 9 trials. Figure 5 is for the same boys as Fig. 2 and Fig. 6 the same as Fig. 3.

When these three figures are superimposed upon one another it is found that not even at the beginning are there any marked differences. The same general tendency appears as in the star tracing curves—a rapid early improvement, with less after the first few trials and much crossing of lines.

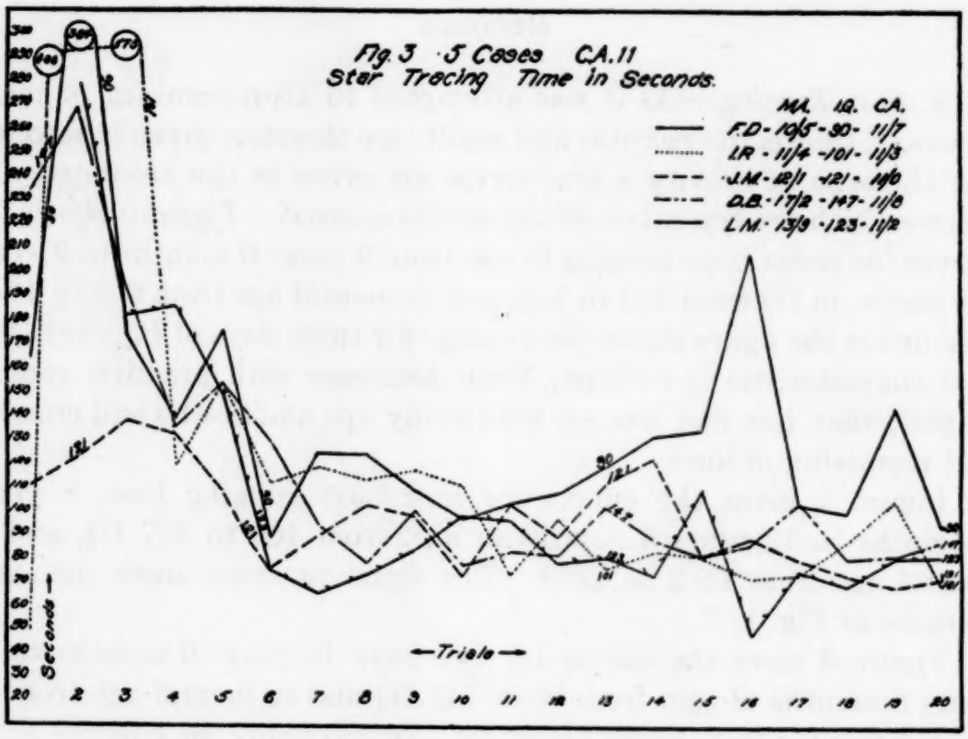


Fig. 3 5 Cases C.A. 11
Star Tracing Time in Seconds.
MA. IQ. CA.
F.D. - 10/5 - 90 - 11/7
L.R. - 11/4 - 101 - 11/3
J.M. - 12/1 - 121 - 14/0
O.B. - 17/2 - 147 - 11/8
L.M. - 13/9 - 123 - 11/2
Seconds
Trials

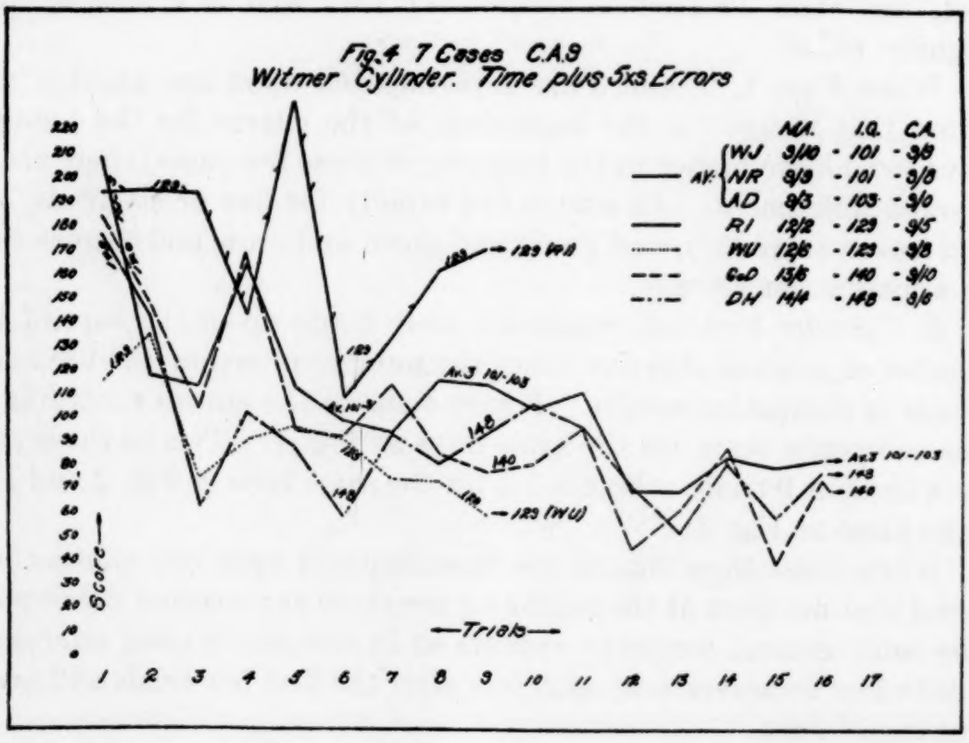


Fig. 4 7 Cases C.A. 9
Witmer Cylinder. Time plus 5xs Errors
MA. IQ. CA.
Av. WJ 9/10 - 101 - 9/8
NR 9/9 - 101 - 9/8
AD 9/3 - 103 - 9/0
RI 12/2 - 129 - 9/5
WU 12/6 - 123 - 9/8
GD 13/6 - 140 - 9/10
DH 14/4 - 148 - 9/6
Score
Trials

Greatly extended experiment may reveal characteristics belonging to groups of boys of varying ages or IQ levels in the behavior of their learning curves in these two tests. In the data gathered there is little to show that either superior IQ or superior experiences as measured by chronological age in any way contributed to higher scores. Not even the discovery and use by the brighter boys of the sequences in the cylinder holes (which were made by most of them quite early) showed any effect in better time or fewer errors. Possibly the method of scoring is at fault, but the writer sees no reason in the method used which would tend to discount the factor of mental ability.

One suggestion that may be fruitful in further investigation along this line came during the experiments and can be made clear, perhaps, by referring to Figs. 1 and 4. In this group there are two boys, W. U. and R. I., of IQ 129, each, and ages 9–8 and 9–5, respectively. In Fig. 1 for star tracing, the curve of W. U., although starting in third place, gets to first place on trial four and keeps that place for the rest of the experiment. On the other hand B. I. made a miserable record for the first three trials, but finally came up to the others, his curve crossing the others, except W. U.'s several times before his last trial, the seventeenth. In fact his time for the experiments was so much greater than anticipated that we got far behind schedule and had to conclude his tracings with the seventeenth for lack of time. In Fig. 4 for the cylinder his low score puts his curve distinctly away from the others while that of W. U. is again plainly the most consistently best score. Unfortunately these two boys had only 9 trials with the cylinder, the number originally planned but after a few experiments increased to 16. The significant thing in connection with these two cases is, of course, the consistently wide difference between them. The suggestion which came in this connection is that temperamental differences had much to do with the differences between their scores. B. I. was talkative, self-appreciative, given to discussing irrelevant matters, fond of his own humor, physically restless, apparently quite unconcerned, except now and then, about his tasks. On the other hand W. U. was quiet, attentive, had exceptional mental and motor self control and apparently took great pains to do his work well.

Criticism is made that the failure of the time curves in the star tracing to discriminate bright and normal IQ subjects is due to the variability in accuracy of tracing, which error curves might reveal. In answer, it is admitted that there is variability in accuracy results, but this very fact evidences native differences under controlled condi-

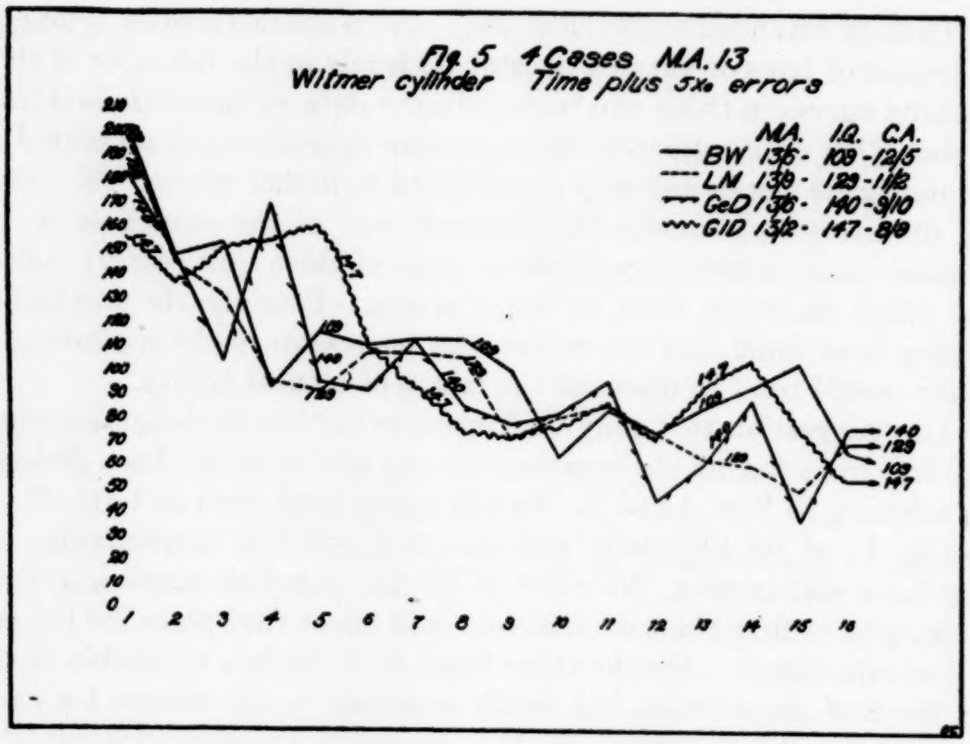


Fig. 5 4 Cases M.A. 13
Witmer cylinder Time plus 5x errors
M.A. I.Q. C.A.
BW 13/5 - 109 - 12/5
LM 13/9 - 129 - 11/2
GeD 13/6 - 140 - 9/10
GID 13/2 - 147 - 8/9
140
129
109
147
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

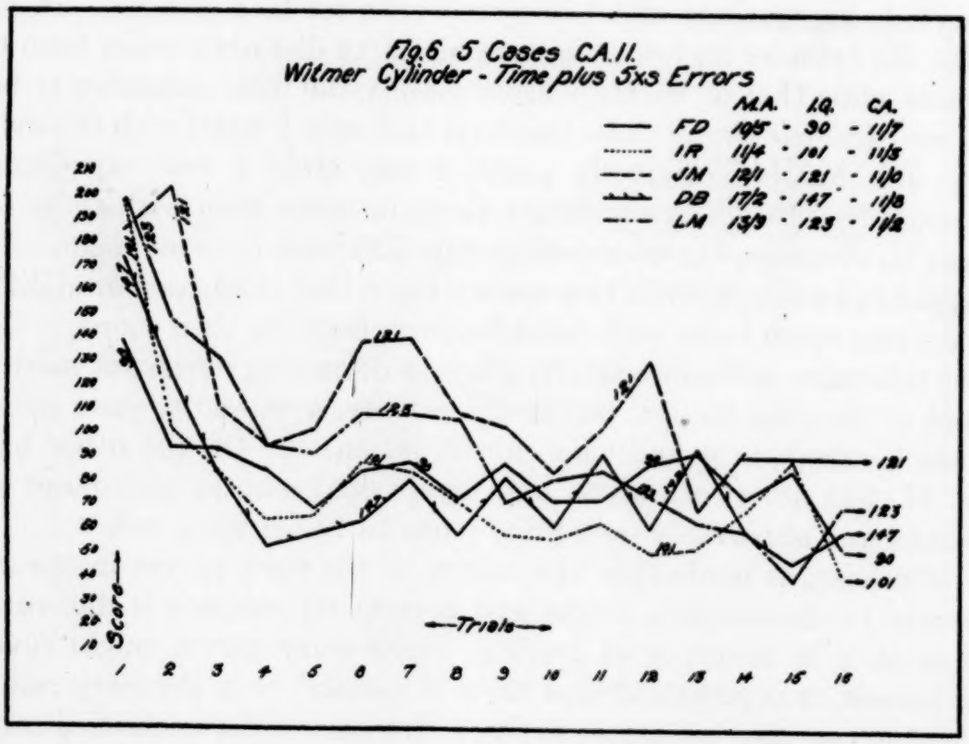


Fig. 6 5 Cases C.A. 11.
Witmer Cylinder - Time plus 5xs Errors
M.A. I.Q. C.A.
FD 10/5 - 90 - 11/7
IR 11/4 - 101 - 11/5
JM 12/1 - 121 - 11/0
DB 17/2 - 147 - 11/8
LM 13/9 - 123 - 11/2
121
123
147
90
101
Score
Trials
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

tions, which were constant cautioning—as necessary—to keep on the line, apparatus, etc. Suppose the impossible feat of making all conditions outside the nervous systems of subjects absolutely uniform. There still would be, by all evidences of past experimentation, variability in results. It may be that, unaware, the experimenter disregarded quality and controlled the star tracing so as to secure the quite uniform time scores. Stranger things have happened. However, first it was anticipated that considerable differences in time rather than uniformity would be found, especially between the highest and lowest IQ subjects. Second, the writer can discover no way in which control would produce the very same relations between high and low IQ in the cylinder test as scored and administered as the graphs show to have been the case. The latter test, therefore, may be considered a check on or criterion of the reliability of the star tracing results. Other explanations than the bearing of accuracy on speed in the star tracing are necessary, and explanations that will account for failure to discriminate between high and average IQ boys in learning both star tracing and the cylinder tasks.

A EARSON-*r* FORM FOR USE WITH CALCULATING MACHINES[1]

LUTON ACKERSON

Institute for Juvenile Research, Chicago

This form for the computation of Pearson product moment correlation coefficients has been devised for statistical laboratories equipped with the usual calculating and adding machinery. The fixing of the "assumed mean" at zero permits the eliminating of deviations and product moments of negative sign which renders the computation work more simple and straightforward and at the same time avoids an annoying source of arithmetical errors, and also permits the printing of the numbers representing the coded step-intervals in the cells themselves which facilitates the plotting of the paired measures in drawing up the scatter-diagram. This preliminary "coding" of the raw measures on the original data sheet as described by Toops (*Journal of Applied Psychology*, 1922, Vol. VI, pp. 172–184) is especially advisable in the case of computing intercorrelations of several variates when each series of measures must be read off several times. If the problem calls for only the correlations of a criterion against a number of variates without other intercorrelating it is still advisable to set the criterion-measures as the X-variate and code them in advance into 15 to 20 or more step-intervals. For cases in which a preliminary coding is not feasible, as in drawing up scatter-diagrams from data in ungrouped form by the use of Hollerith or Powers card-sorting machines, one may be guided by grouping-limits written into the spaces along the top and side of the grill and disregard, of course, the code numbering within the cells.

The fixing of the assumed mean at zero requires that larger numbers must be handled in computations than with the use of other printed Pearson-*r* forms which permit setting the assumed mean nearer the actual mean. Therefore the present form may not possess any advantages over other forms if the computation must be done "by hand" except in working upon small populations in which the arithmetical labor is not onerous.

---

[1] For many helpful suggestions in devising this form the writer is indebted to Professor L. L. Thurstone, and to Miss Margaret O'Connor, statistical clerk, and Miss Marion Mason, artist, of the Institute for Juvenile Research.

The computation work is very simple and can be done adequately by any clerk with a fair knowledge of high school algebra. The entries under column f_y are the *frequencies* of the Y-scores (reduced or unreduced) for each horizontal Y-array, and the entries under row f_x are similarly the *frequencies* of the X-scores (reduced or unreduced) for each vertical X-array. The entries under columns f_yY and f_yY^2 are sufficiently indicated as the products of the entries in the f_y and Y columns, and in the f_yY and Y columns, respectively. The entries under rows f_xX and f_xX^2 are similarly obtained from the f_x and X rows.

Each entry under column ΣX_y (the sum of the X's for each Y-array) is the *sum* of the products of each cell-frequency by its step-interval code number for the given horizontal Y-array. The entries under column $Y \cdot \Sigma X_y$ are the products of the entries in the ΣX_y and Y columns. This simplified method of calculating the troublesome XY-products has been suggested by Thurstone, to whose article (*Journal of Educational Research*, 1922, Vol. VI, pp. 49–56) the reader is referred who wishes a more detailed description of the process. By using a calculating machine equipped with a "carry-over dial" such as a Marchant or Monroe, one may obtain the entries for columns f_y and ΣX_y in one series of operations by setting the successive X printed code numbers in the keyboard and allowing the sum of the frequencies to accumulate in the multiplier dial.

The total sum of column ΣX_y is equal to the total sum of row f_xX. This equality serves as an adequate check for the arithmetical accuracy of the entries in columns f_y and ΣX_y and rows f_x and f_xX. A check for the arithmetical accuracy of the remaining entries in columns f_yY, f_yY^2, $Y \cdot \Sigma X_y$ and row f_xX^2 may be made directly from these verified entries in the extension columns and rows without the necessity of reviewing the cell-frequencies themselves.

The six quantities to be supplied in the r-formula are the summations indicated in the lower right corner. In the formula for the standard deviations, the values of $\sqrt{N \cdot \Sigma X^2 - (\Sigma X)^2}$ and $\sqrt{N \cdot \Sigma Y^2 - (\Sigma Y)^2}$ are already obtained in the solution of the r-formula above.

The form is designed to cover a sheet 17 by 22 inches which may be easily folded into convenient letter-size for filing. In this size "planographed" copies can be made up at a cost of about 25 dollars for 1000 or about 33 dollars for 2000 copies.

(This is one of a series of articles published by members of the staff of the Illinois Institute for Juvenile Research, Chicago, Dr. Herman M. Adler, Director, Series C, No. 124.)

A SIMPLE APPARATUS FOR THE CLASS DEMONSTRATION OF THE CONDITIONED RESPONSE

GEORGE HUMPHREY

Queen's University, Kingston, Ontario

The apparatus to be described may be used to show the essential properties of the conditioned response to a group of as many as 20 students, or possibly more with special arrangements. It is simple to make and operate. Excluding the stimulus apparatus and the batteries, it may be put together for less than five dollars.

Essentially it consists of a modification of the Watson apparatus, with the front electrode, on which, in Watson's scheme the finger rests, movable about a hinge. This electrode rests on a spiral wire spring which, when the subject's hand is lifted, forces the brass plate upwards to an angle of about 30 degrees with the base board. The electrode is prevented from making a wider excursion by contact with the lower of a pair of strips of brass fastened together to form a narrow V, with platinum contacts on the inner sides of the open ends. When the electrode is thus forced upwards, the V is closed and the platinum points meet, closing the circuit of a glow lamp at the operator's table. When the subject's hand is placed on the hinged plate, this is flattened to the base board by the weight of the hand. This allows the V to expand by the tension of the brass strips composing it, and the platinum points separate, thus breaking the circuit and extinguishing the glow lamp. When the shock is given and the subject raises his hand, the legs of the V are forced together and the platinum points meet, making the circuit again. Thus the lamp indicates to the operator whether the subject's hand is on or off the electrode, when the light is shining the hand being off, and *vice versa*. This serves two purposes. When the light is off and a stimulus given, the lighting of the lamp indicates to the operator behind the screen that the subject has responded by lifting up his hand. When the light goes off again this is a signal that the subject is ready for the next stimulation.

For the electric stimulation (primary stimulus) we have found most satisfactory two two-volt storage batteries in conjunction with a Harvard Porter inductorium, set at from two to five, the setting of the latter being varied with the age and strength of the subject. Contact for this stimulation is most easily and silently made by pressing the bare wire on a brass plate screwed to the table, this method being

found preferable to a telegraph key or other device. It is important, when one is working with arpeggios or scales or other more complicated forms of auditory stimulation, to ensure the minimum amount of movement, and thus distraction, on the part of the operator. For the secondary ("conditioned") stimulus we have used a metal instrument of the Xylophone type,[1] one with cylinders being preferable on account of the relatively pure tone it gives. With a little practice scales, arpeggios, and simple melodies can be played on it not unpleasingly.

When the apparatus is used for class demonstration, the operator, who may be a member of the class, sits with his back to the group and facing the apparatus. With his right hand he holds the hammer for the instrument, with his left the wire for the electric stimulation. The subject is screened from the class[2] which, watching the lamp, records the response. The series of stimulations to be given is prearranged, so that all that is necessary is to mark + or −, according as the light does or does not go on. The conditioned response will appear in about 30 trials with most subjects, though up to 300 or 400 trials are necessary before it becomes well established. It is important that at the beginning of each session the subject should have read to him a cautionary direction, telling him that he is not undergoing a test of any kind but is to act entirely naturally.[3]

With the set up described, there may be demonstrated the following facts reported by Pavlov.

1. The formation of conditioned responses.
2. The discrimination of the same responses to a single tone.
3. The early generalization of the response whereby any like stimulus will call forth the response before discrimination.

---

[1] Our apparatus had the trade name of Zellophone, and was furnished locally through the Nerlich Company of Toronto. It is imported from Germany, and has the advantage that it is a real musical instrument and not a toy. It may be necessary to match up a set of cylinders in tune from two instruments if arpeggios and melodies are to be played. The cost in Canada for an instrument with 12 cylinders is two dollars.

[2] He may preferably be behind a partition, where he can hear the instrument but not see the class.

[3] For the standard form used by us, see an article by the writer in the *Journal of Abnormal and Social Psychology*, July, 1927, entitled: The Effect of Sequences of Indifferent Stimuli on a Reaction of the Conditioned Response Type. This article describes also a more elaborate form of the same apparatus.

4. The temporary disappearance of the established response by repetition of the secondary stimulus without the primary one.

5. The permanent disappearance of the response when the secondary stimulus is repeatedly given unsupported by the primary one.

6. The inhibition of the response by repeatedly sounding a note which is near to the conditioned stimulus but which has already been discriminated out. (This was one of the facts that led Pavlov to formulate his doctrine that discrimination is accomplished by inhibition. It is best exhibited when the discrimination is not quite completed, when it will be found that to train out the "wrong" note will often be followed by the disappearance of the response to the "right" note.[1]

7. The inhibition of the response by irrelevant stimuli in general. This may best be shown by including the active note in a sequence composed of non-active notes.[2]

The disadvantages of the apparatus are that it shows only the fact, not the time or the extent of the response, and that minimal or negative responses are not indicated. It also, of course, gives no permanent record. In conclusion, it may be noticed that we have had most success with paid subjects over 14 years old, who should first be given the Seashore test for acuity of pitch discrimination. We spent many fruitless hours with one boy who later turned out to be very seriously deficient on this test.

CORRECTION

Due to a delay in receiving proof of the article, "A Test of Scientific Aptitude," by D. L. Zyve, which appeared in November, 1927, journal, several changes desired by the author were not made. On page 533, the seventh line from the top should read: (The correct answer is 21). On page 539, the sixth line from the bottom should read: $r_{\text{TSA-Sch.}} = .019 \pm .091$. The third line from the bottom of same page should read: $r_{\text{Th-Sch.}} = .41 \pm .07$.

---

[1] Humphrey George: *Journal of Abnormal and Social Psychology*, July, 1927.

[2] See the same article for discussion of this phenomenon, with possible alternative explanations.

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES

REPORTED BY JAMES E. MENDENHALL
Research Associate, Institute of Educational Research, Teachers College
Columbia University

Intelligence Testing

An Investigation into the Comparative Intelligence and Attainments of Jewish and Non-Jewish School Children. The British Journal of Psychology, October, 1927, 135–146. 1894 Jew and non-Jew elementary school children of London were tested. On the Northumberland Standardized Tests (General Intelligence, English, and Arithmetic) Jewish children proved definitely superior to the non-Jewish children attending the same schools, the superiority being more marked with the boys than with the girls.

A Study of One Hundred Non-promoted Children. Elizabeth M. Stalmaker and R. D. Roller, Jr. Journal of Educational Research, Nov., 1927, 265–270. The group ranging from Grades I to VI were tested with the Stanford-Binet and the Porteus Tests for Mental Deficiency. "The failure of probably 91 per cent and certainly of 31 per cent was due to mental deficiency . . . The Stanford Binet test measured more nearly the abilities demanded for success in school subjects than did the Porteus test."

Psychological Testing in a Women's College. Margaret R. Davidson and Andrew H. McPhail. The Personnel Journal, Dec., 1927, 266–275. This study reports the results of four years' testing with the Brown University Psychological Examination. Data as presented tends to show (1) correlations between psychological test and scholastic record for one semester range from .47 to .57, (2) striking differences in academic course grades between two highest tenths and two lowest tenths in psychological test, (3) the number of drop outs of the two lowest tenths to be much greater than that of the two upper tenths, (4) that rank in the preparatory senior class should be also used for predictive purposes.

Psychology of School Subjects

The Additive vs. the Take-away Method of Teaching Subtraction Facts. B. R. Buckingham. Educational Research Bulletin, Sept. 28, 1927, 265–269. Small equivalent groups of children who had no previous practice in substraction were practiced in the additive method, and in the "take away" method. The "take away" method of subtraction proved superior for immediate retention.

How Shall We Subtract? W. J. Osburn. Journal of Educational Research, Nov., 1927, 237–246. A subtraction test was given to 1414 pupils in the fourth, fifth and sixth grades of 23 school systems. Record was made by each pupil of the subtraction method used. The group using "take away equal additions" had the least average number of errors; the "take away decomposition" group the next

largest average number of errors; and the "additive" group the largest average number of errors. The "take away decomposition" method is in most common usage in elementary schools.

Bilingualism (with special reference to Bengal). Michael West. Bureau of Education, India, Occasional Reports, No. 13, 1926. An extensive and intensive study of the problem of bilingualism as it relates to methods of teaching reading and other school subjects. Ingenious objective techniques have been applied to determine the optional devices for learning.

Methods of Study

Testing and Training the Inferior or Doubtful Freshman. Edward S. Jones. The Personal Journal, Oct., 1927, 182–191. The use of drill in writing, rapid reading, note taking, and mathematics problems helped the inferior freshman to raise the quality of his work.

Study Methods of College Students in Relation to Intelligence and Achievement. Clay C. Ross and Nira M. Klise. Educational Administration and Supervision. Nov., 1927, 551–562. A questionnaire reveals no significant differences in methods of study between the successful group and the unsuccessful group of students. However, the writers "are of the opinion that success comes only through *a good method skilfully used.*" Also that "of those students who have had specific instructions in how to study the least intelligent have profited most."

How Well College Students Read. William F. Book. School and Society, Aug. 20, 1927, 342–348. The reading ability of college students was measured. This was followed by classes in "How to Study," which the author found very helpful to students in improving reading ability.

Vocational and Educational Guidance

Determination of Vocational Aptitudes. Does the Tapping Test Measure Aptitude as Typist or Pianist? Harry D. Ketson. The Personnel Journal, Oct., 1927, 192–198. Comparisons of average tapping rates were made between an unselected group of 25 and (*a*) 25 pianists and (*b*) 25 typists at the University of Indiana. Differences (not statistically significant) were found between the groups. The author states the hypothesis "that 50 per cent of the population can succeed with a 50 percentile degree of success in 50 per cent of the occupations."

Vocational History of Psychologists. Harry D. Ketson. The Personnel Journal, Dec., 1927, 276–282. Data regarding the male members of the American Psychological Association revealed the median ages at which advancement was secured.

Character and Personality

The Reliability of Self-scored Measures. Lloyd N. Yepsen. School and Society, Nov. 19, 1927, 657–660. In classes in mental measurement, 53 students were tested with a literacy test. Each individual was allowed to score his own paper. Two experiments showed that 24.5 and 29.5 per cent respectively of the papers were changed during correction by self. "We cannot advise, however, the practice of allowing *any* individuals, be they school children or adults, to correct their own papers."

Resourcefulness, an Unmeasured Ability. Robert O. Beauchamp. School Science and Mathematics, May, 1927, 457–465. Resourcefulness (as displayed by 64 eleventh and twelfth grade high school students) in solving chemistry experiments in the laboratory and then in writing was found to be low by correlation with intelligence as measured on the Otis Group Intelligence Test and the McCall Multi-mental Test.

Adaptation Difficulties in College Students. Karl A. Menninger. Mental Hygiene, July, 1927, 519–535. An attempt was made to differentiate between a "healthy" minded group and an "unhealthy" minded group of college students by their responses to a questionnaire.

British Norms for the Pressey Cross-out Test. Mary Collins. The British Journal of Psychology, Oct., 1927, 121–133. The test was given to over 1500 children in Scotland and England, with ages ranging from 11 to 15. Retests showed a great variability from original responses by the same pupil. Some of the most common responses of English children differed from the American norms. "The test as a test of any aspect of character, is not reliable . . . Its value lies in the diagnostic field, as a guide to the emotional disturbance."

Some Factors Affecting Achievement

The Relation of Achievement between College Fraternity and Non-fraternity Groups. Alvin C. Eurich. School and Society, Nov. 12, 1927, 624, 630. The registrar's records for 2,817 students including 36,174 marks for a period of 11 years of the University of Maine were used in this study. On comparing fraternity and non-fraternity records the results "suggest that a fraternity environment does not affect the scholastic achievement of the average college student," when either graduate or non-graduate groups are considered.

Progress of Equivalent One-room and Graded-school Pupils. C. W. Stone and J. W. Curtis. Journal of Educational Research, Nov., 1927, 260–264. 420 pupils in the seventh, eighth and ninth grades were tested with a battery of mental and educational tests. 114 pupils paired in intelligence were secured. The results indicated that the graded school pupils were by test norms 4 to 5½ months superior educationally to the one-room school pupils.

The High School Achievement of Children Trained under the Individual Technique. Carleton Washburne and Louis Edward Raths. The Elementary School Journal, Nov., 1927, 214–224. The 323 Winnetka high school pupils who had had individual instruction in an elementary school proved superior in scholastic achievement and in extra-curricular participation for other pupils in the same township high school.

Measurement and Analyses of Student Achievement in a Beginning Course in Educational Psychology. Winona M. Perry. Education, Sept., 1927, 12–27. The justifications of answers to true-false statements are required in order to show whether students comprehend the statements.

Miscellaneous

Analyses of Three Thousand Illegibilities in the Handwriting of Children and of Adults. Luella C. Pressey and Sydney L. Pressey. Educational Research Bulletin, Sept. 28, 1927, 270–273. The papers of 650 children and adults in the

third to college grades were examined. Characteristic errors in handwriting were thus determined.

Possibilities of Tests on Physical Education. David K. Brace. American Physical Education Review, Sept., 1927, 506–513. Oct., 1927, 563–568. A suggestive paper outlining the possibilities of developing standard measures of progress in physical abilities. A number of usable tests in this field are presented.

New Tests

Arithmetic Neatness Scale. F. B. Knight and W. E. McClure. Scott, Foresman and Company, 1927. "Samples were collected not only from typical classroom situations but from men and women in many walks of life."

Langren-Woody Reading Test. Paul V. Langren and Clifford Woody. World Book Company, Yonkers, N. Y., 1927. A test designed to measure word meaning, rate, fact material, total meaning, central thought, following directions and organization for Grades 4*B* to 8*A*.

Columbia Research Bureau American History Test (for High Schools and Colleges—two forms). Harry J. Carman, Thomas N. Barrows, and Ben D. Wood. The World Book Company, Yonkers, N. Y., 1927. This test covers the historical aspects of political, social, economic, religious, and educational questions. Knowledge of facts, reasoning ability and discriminating judgment are involved in answering many of the test items.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

The Importance of Word Meaning in the Construction of Reading Books

Reading and Word Meanings, by Edward William Dolch, Boston: Ginn and Company, 1927. Pp. IV + 129.

This is a book which ought to help shift the attention and emphasis in reading from the acquisition of word forms to an understanding of word meanings. It is an important analysis of the "meaning" content of school reading books. Mr. Dolch's introductory statement reveals the central purpose of the book: "If reading is to be considered a thought-getting process then for the purpose of studies in reading the term 'word' must refer to an idea conveyed by a symbol and not to the symbol itself; it must refer to a meaning and not to a spelling." That sounds rather obvious, yet strangely enough it has scarcely been considered in previous studies of the content of readers. No one can deny, however, that it is the really fundamental and sound aspect of the reading process.

A study of word meanings in a single set of school "readers" was conducted by the author. His investigation was unusual because previous studies had been made of readers of the same grade. Mr. Dolch's analysis disclosed, for example, that each of the 70 spellings in the first reader had, on the average in the five grades, from two to six meanings. In the set of five readers there were 536 spellings; these 536 separate spellings, however, were used with 1358 meanings, that is, each spelling was used on the average with three different meanings.

This analysis suggested a number of important problems. For example, how many meanings should be used in a given grade? How often should each meaning be repeated? The account shows that there were 5455 meanings in the entire set of readers. This number would require about seven new words per reading lesson in the fourth and fifth grades. Tabulation of frequency of repetition showed that

many words were repeated in the first grade. In Grade II, however, one-third of the new meanings were used only once. This practice of not repeating meanings was revealed even more strongly in Grades III, IV, and V, in each of which two-thirds of the new meanings were used but once.

The problem of difficulty of words was also studied. A test was given, using the meanings found in the analysis. Children were asked to help find the hard words in the readers. The test was not one of achievement—to get a score—it was merely to check words previously "studied." The results were compared with other available word lists. Interestingly enough, there was positive correlation between them, but it was low. Indeed the author's results throw grave doubt upon the validity of existing word lists.

Discussion of the procedure used in this study and of the problem suggested by the approach occupy the greater part of Mr. Dolch's book. In my judgment it is very significant. True, it is only a beginning and the results are not to be regarded as ultimate data. The study, however, opens up a whole comparatively new field for scientific investigation. For example, how do children get new meanings? How many new meanings can they acquire in a day? How much repetition should there be? What should be the distribution of meanings? Which word should be included? Which words are difficult? What kinds of tests for meaning should be used?

These and other questions have been suggested and discussed in Mr. Dolch's monograph. In addition, there is a helpful chapter on the editing of reading books to insure meaningful understanding. Valuable suggestions to authors and publishers are also included.

LOUISE C. KRUEGER.

Browning School, New York City.

NEW DEVICES FOR PRIMARY READING

An Experimental Study of the Beginning Reading of Deaf-mutes, by Helen Thompson. New York City: Teachers College, Columbia University Contributions to Education, No. 254, Bureau of Publications, 1927. Pp. IV + 81.

Dr. Gates has advisedly pointed out in a recent issue of the Journal of Educational Research, the need for training the deaf-mute in effective reading habits. He urged that the time now devoted to teach-

ing the mute to talk and read orally would better be spent in teaching the mute to read silently. This study undertook the task of constructing materials suitable to the instruction of the deaf in reading, and also of experimenting with the materials. The results were checked against those which were secured from an equivalent control group—each of the two groups being composed of ten deaf-mute children.

The writer not only tested the groups from time to time, but also kept a very careful day to day record of the method used and material taught each day and the reactions of the children. The procedure followed in this "planned in advance" presentation of materials of reading is commendable. The reading material is so set up as to take into consideration the important psychological principles of learning—one thing is introduced at a time and learned thoroughly, comparison and contrast are used, repetition of words is provided for, pupil activity is encouraged.

Aside from the excellence of the reading material invented by the author, the difficulties encountered by these mute children are exceedingly suggestive as to the difficulties which hearing children may be facing. The average hearing child with his several avenues of learning masters reading eventually despite the methods by which that reading is taught. But the waste of time, effort, and interest due to such desultory methods of teaching reading should be a cause for concern. It is evident that the deaf-mute child is almost entirely dependent upon the sense of sight. This limitation makes it exceedingly difficult to teach some words such as "can," "where," and "under." Normal children in all likelihood are confronted with similar troubles. The light which this study throws upon the psychology of beginning reading is probably its most significant contribution.

This book contains samples illustrative of the material used in the experiment. Much ingenuity was exercised in constructing the reading matter so that the many learning factors would be considered. The multitude and variety of sound psychological devices if used in the future should go far toward aiding any child in the rapid and sure acquisition of meanings through reading. Anyone interested in the problem of reading will do well to examine this material with great care.

JAMES E. MENDENHALL.

The Lincoln School, Teachers College.

An Objective Attack upon University Problems

Research Adventures in University Teaching, by S. L. Pressey, D. A. Worcester, J. M. Ferguson, J. A. Seaton, L. C. Pressey, O. R. Chambers, H. J. Arnold, H. C. Martin. Bloomington, Illinois: Public School Publishing Company, 1927. Pp. XVIII + 151.

This book is a collection of studies regarding college problems in the fields of study, curricular adjustment, emotional and characterological development, previous preparation, and teaching. The studies have been narrowly defined in their scope, which in turn has added practicability to the findings.

The major problem which faces the university today is that of adapting the school to the needs of the individual student. This adaptation has been proceeding apace for some little time; however, many of the past attempts to achieve this end might well be characterized as sporadic, of limited scope, and highly subjective. A book incorporating findings of a comprehensive, definite, and objective nature is indeed to be welcomed.

The authors have given attention to the many phases of the student's life; they have undertaken to find out the deficiencies of incoming groups; they have, moreover, subjected some of the theories regarding university teaching to a test. The methods by which students can be helped in preparing their work, and some economical methods of conducting college courses—both as presented by the writers are worthy of note.

Above all, the group of authors should be commended on the effort to *cooperatively* recreate university life by scientific devices. Out of the researches presented in this little book should grow a host of worthwhile studies in the college field. The writers have indeed followed a good policy in writing down in explicit and organized form the results of their several investigations.

This is a book deserving of careful study by all those engaged or about to engage in university or college teaching.

James E. Mendenhall.

The Lincoln School, Teachers College.

Understanding the Inhuman in Human Nature

Understanding Human Nature, by Alfred Adler. (Translated from the German by W. Beran Wolfe, M.D.) New York: Greenberg, Publisher, 1927. Pp. XIII + 266.

Another explanation of "man's inhumanity to man" is offered in this translation of a series of popular lectures given in Vienna by the psychoanalyst, Dr. Alfred Adler. It does not purport to be an exhaustive treatment of the problems of psychopathology, but is merely "a manual of suggestions and stimulation," "an attempt to acquaint the general public with the fundamentals of Individual Psychology" (which is the term Adler uses to characterize his system of psychotherapeutics).

Whereas it will undoubtedly be illuminating to teachers, parents, and just "ordinary people," the students of objective methods will find much to quarrel with in Dr. Adler's subjective treatment of attitude and behavior, with his postulate of "soul" and with his acceptance of the "logic of communal life as though it were an ultimate absolute truth." He passes lightly over the minute analyses of structure and function with which psychology usually begins and often ends. Instead he directs attention at once to a practical revelation of the individual's inner and overt behavior pattern.

He ascribes most of human endeavor to a striving for recognition and security in the group life.

The feeling of inferiority and insecurity is constantly present, he points out, in the consciousness of man. Even the earliest strivings of the infant for the satisfaction of its wants contain this struggle for recognition. Hence, Dr. Adler puts an enormous emphasis upon the early days of infancy and childhood, both as possibilities for right education, and as sources for the explanation of maladjustments in later life. "The examination of the soul life of the child is the fulcrum of our science." Again, "The feeling of inferiority and the striving for recognition is thus fundamental to every act and situation in life, and the methods by which the individual has become accustomed to achieve satisfaction toward these ends mark the outlines of his character."

Whether we agree with his fundamental hypotheses or not, the book does offer a stimulating and intelligible insight into the devious "psychic detours" which the human organism—call it consciousness if you will—makes in order to achieve satisfaction for itself.

Ann Shumaker.

Graduate Student, Teachers College.